

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 84 ELLVGLTLGGYKTGTTSTSVNETDAAGDPMYLTFTSQDGNNHQFTTKVIGKDSRDFDIS 143

```

Db 61 ELIVGLITLGGITGTTSTSVNFTDAGDPMTFTSODSNNHQTFTKVLGKDSRDPDIS 120
QY 144 PKVNGENLVGDVYVLTATGSDFFVRSIGSGKGLAAGKYDVAITYVSNQ---EFMIRAY 200
    |||||
Db 121 PKVNGENLVGDVYVLTATGSDFFVRSIGSGKGLAAGKYDVAITYVSNQSTIEGRIRAY 180
QY 201 EONPQHFIEDLEKVRVQOLGHSVLEELVQVVKDKNDISIKYPRKDSYFANRVIT 260
    |||||
Db 181 EONPQHFIEDLEKVRVQOLGHSVLEELVQVVKDKNDISIKYPRKDSYFANRVIT 240
QY 261 DDELLKLIATVFLPEDTILKGGHYNOLONGIKRVKEFELESSPNTOMELRAFMAYVHS 320
    |||||
Db 241 DDELLKLIATVFLPEDTILKGGHYNOLONGIKRVKEFELESSPNTOMELRAFMAYVHS 300
QY 321 LPAIDRDDILKIVDSMNHGDAKSLREELAEITAEIKTVSVIAEINKHLSSSGTIN 380
    |||||
Db 301 LPAIDRDDILKIVDSMNHGDAKSLREELAEITAEIKTVSVIAEINKHLSSSGTIN 360
QY 381 IHDKSINLMDKNLYGTDEEIFKASAEYKLEKMPOTTIOVDGSEKKIYSIKDFLSENK 440
    |||||
Db 361 IHDKSINLMDKNLYGTDEEIFKASAEYKLEKMPOTTIOVDGSEKKIYSIKDFLSENK 420
QY 441 RGALGNLNKNSYSYNKDNNELSHFATCSDKSRPLNDVYSOKTTOISDITSRENSAIEAL 500
    |||||
Db 421 RGALGNLNKNSYSYNKDNNELSHFATCSDKSRPLNDVYSOKTTOISDITSRENSAIEAL 480
QY 501 NREIQYDYSVMORLDDTSGK 521
    |||||
Db 481 NREIQYDYSVMORLDDTSGK 501

```

RESULT 2

```

ID W01045 standard; Protein: 480 AA.
AC W01045;
DE 28-DEC-1996 (first entry)
DE Y. pestis F1/V antigen fusion.
KW Plaque; vaccine; genetic immunisation; V antigen; lcrV;
KW F1 antigen; calf.
OS Chimeric Yersinia pestis strain GB;.
OS Chimeric synthetic.
FH Key
FT Location/Qualifiers
FT 1..149
FT /label= "F1 antigen"
FT /note= "mature F1 antigen"
FT 150..155
FT /label= "linker"
FT /note= "6-amino acid peptide linker"
FT 156..480
FT /label= "V antigen"
FT /note= "mature V antigen"
FT protein
FT W09628551-A1.
FT 19-SEP-1996.
FT 13-MAR-1996: G00571.
FT 13-MAR-1995: GB-005059.
FT 15-SEP-1995: GB-018946.
FT 05-DEC-1995: GB-024825.
FT (MINA) UK SEC FOR DEFENCE.
FT PA Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED.
FT WPI: 96-43382/43.
FT N-PSDB: T38256.
FT Yersinia pestis V antigen and F1 antigen or their protective
FT epitope parts - useful in vaccine for protection against plague
FT disclosure; Page 51-55; 98pp; English.
FT A fusion protein (W01045) comprises the F1 antigen (see also
FT W01042) and V antigen (see also W01041) of Yersinia pestis joined
FT by a linker that allows each protein to attain its conformational
FT state. It is the product of a gene fusion (T38256) obt. by PCR
FT amplification of Y. pestis DNA. F1/V fusion protein can be
FT expressed by gut-colonising organism transformants, to induce
FT an immune response against Y. pestis, the causative organism of
FT plague.

```

SQ Sequence 480 AA:

Query Match 89.4%; Score 2382.5; DB 1; Length 480;
 Best local similarity 98.5%; Pred. No. 1,1e-163;
 Matches 473; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

```

QY 45 ADITASTATATFALVEPARITLTYKREGAPITIMDNGNIDELLVGLTLGGYKGTSTSV 104
    |||||
Db 1 ADITASTATATFALVEPARITLTYKREGAPITIMDNGNIDELLVGLTLGGYKGTSTSV 60
QY 105 NPTDAGDPMYLTFTSODSNNHQTFTKVLGKDSRDPDISPKVNGENLVGDVYVLTATGSD 164
    |||||
Db 61 NPTDAGDPMYLTFTSODSNNHQTFTKVLGKDSRDPDISPKVNGENLVGDVYVLTATGSD 120
QY 165 FVFRSISGSGKGLAAGKYDVAITYVSNQ---EFMIRAYEONPQHFIEDLEKVRVQOLTG 221
    |||||
Db 121 FVFRSISGSGKGLAAGKYDVAITYVSNQSTIEGRIRAYEONPQHFIEDLEKVRVQOLTG 180
QY 222 HGSSVLEELVQVVKDKNDISIKYPRKDSYFANRVITDIELLKLILAYFLPEDTILK 281
    |||||
Db 181 HGSSVLEELVQVVKDKNDISIKYPRKDSYFANRVITDIELLKLILAYFLPEDTILK 240
QY 282 GGHYNQLONGIKRVKEFELESSPNTOMELRAFMAYVHSLPADRIDDDILKIVDSMNH 341
    |||||
Db 241 GGHYNQLONGIKRVKEFELESSPNTOMELRAFMAYVHSLPADRIDDDILKIVDSMNH 300
QY 342 GPASRLREELAEITAEIKTVSVIAEINKHLSSSGTINIHDKSINLMDKNLYGTDEI 401
    |||||
Db 301 GPASRLREELAEITAEIKTVSVIAEINKHLSSSGTINIHDKSINLMDKNLYGTDEI 360
QY 402 FRASAEYKLEKMPOTTIOVDGSEKKIYSIKDFLSENKRGALGNLNKNSYSYNKDNNEL 461
    |||||
Db 361 FRASAEYKLEKMPOTTIOVDGSEKKIYSIKDFLSENKRGALGNLNKNSYSYNKDNNEL 420
QY 462 SHFATCSDKSRPLNDVYSOKTTOISDITSRENSAIEALNREIQYDYSVMORLDDTSGK 521
    |||||
Db 421 SHFATCSDKSRPLNDVYSOKTTOISDITSRENSAIEALNREIQYDYSVMORLDDTSGK 480

```

RESULT 3

```

ID R79961 standard; Protein: 329 AA.
AC R79961;
DE 18-APR-1996 (first entry)
DE Partial lcrV (V antigen) of Y. pestis.
KW lcrV; V antigen; virulence; plaque; vaccine; epitope.
OS Yersinia pestis.
OS W095244/5-A1.
PN 14-SEP-1995.
PF 06-MAR-1995: G00481.
PF 08-MAR-1994: GB-004577.
PF (MINA) UK SEC FOR DEFENCE.
PF PI Leary SEC, Titball RW, Williamson ED, Leary SE.
PF WPI: 95-328268/42.
PF N-PSDB: T04222.
PF Recombinant DNA expressing Yersinia pestis V antigen - useful in
PF oral or parenteral vaccines for protection against plague
PF Claim 6; Page 11-13; 25pp; English.
PS R79961-62 are encoded by 104222-23 (resp.), DNA sequences (lcrV) encoding
CC all or a protective epitope part of the mature V protein of Yersinia
CC pestis. The protein was expressed as a fusion protein with maltose
CC binding protein or glutathione-S-transferase in 3 different plasmid
CC vectors. Y. pestis is the highly virulent causative organism of plague
CC in a wide range of animals, including man. The V antigen (lcrV) is an
CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
CC The V antigen is postulated to act as a virulence antigen, and
CC transformed microorganisms contg. recombinant DNA encoding a V antigen
CC protein/peptide are useful in vaccines to protect against plague.
CC Sequence 329 AA;

```

Query Match

61.6%; Score 1641.5; DB 1; Length 329;

Best Local Similarity	99.4%	Pred. No. 1.2e-110;	Matches 326;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;
QY 194	EFMIRAYEQNHQHFEDLEKRVBEQDQJTGSSVLEELVOLVYDKNIDISIKYDPKRDSEV	253					
Db 3	EF-IRAYQNHQHFEDLEKRVBEQDQJTGSSVLEELVOLVYDKNIDISIKYDPKRDSEV	61					
QY 254	FANRYITDIDIELKKIILAYFLPEDTILKGHYDNLQNGIKRVKFELESSPNTOWELRAE	313					
Db 62	FANRYITDIDIELKKIILAYFLPEDTILKGHYDNLQNGIKRVKFELESSPNTOWELRAE	121					
QY 314	MAVNHFSLTARIDDDILKAVYDSNNHHGDARSKRREELAEITALKITYSVQAEINHL	373					
Db 122	MAVNHFSLTARIDDDILKAVYDSNNHHGDARSKRREELAEITALKITYSVQAEINHL	181					
QY 374	SSSGTINIHDSINIMDKNLGYTDEEIEFKASAEYKILEKMPQTTIQYDGSSEKKIVSIND	433					
Db 182	SSSGTINIHDSINIMDKNLGYTDEEIEFKASAEYKILEKMPQTTIQYDGSSEKKIVSIND	241					
QY 434	FLGSENRKTGALGNLKNYSYINKDNNELSHEFATTCSDKSRPLNDLVSQRTQLSDITSRF	493					
Db 242	FLGSENRKTGALGNLKNYSYINKDNNELSHEFATTCSDKSRPLNDLVSQRTQLSDITSRF	301					
QY 494	NSAIEALNRFIOKYDVSVMORLLDDPSGR 521						
Db 302	NSAIEALNRFIOKYDVSVMORLLDDPSGR 329						
RESULT 4							
ID W01040	standard; Protein: 329 AA.						
AC W01040;							
DT 28-DEC-1996	(first entry)						
DE Y. pestis V antigen.							
KM Plague vaccine; genetic immunisation; V antigen; lcrV;							
KW Fl antigen.							
OS Yersinia pestis strain GB.							
FT Key	Location/Qualifiers						
FT Peptide	1..4						
FT	/note="vector-encoded peptide"						
PN W0628551-A1.							
PD 19-SEP-1996.							
PF 13-MAR-1986;	G00571.						
PR 13-MAR-1985;	GB-003059.						
PR 15-SEP-1995;	GB-018946.						
PR 05-DEC-1995;	GB-024825.						
PA (MINA) UK SEC FOR DEFENCE.							
PI Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;							
DR WPJ: 96-433824/43.							
DR N-PSDB: T38242.							
FT Yersinia pestis V antigen and Fl antigen or their protective							
PT epitopic parts - useful in vaccine for protection against plague							
PS Disclosure: Page 25-28; 98PP: English.							
CC Yersinia pestis V antigen (W01040) is capable of evoking protective							
CC immune responses in animals. A combined vaccine of V and Fl							
CC antigens (see also W01042) can at least match the protection							
CC afforded by live attenuated EV76 vaccine without any of the hazards							
CC that have kept the EV vaccine from general use. The V antigen is							
CC produced using the lcrV gene (see also T38242) obtd. from Y.							
CC pests by PCR amplification. It can also be prepd. as a fusion with							
CC Fl antigen (see also W01044-45) and expressed by attenuated Arca or							
CC C Salmonella typhi as a live vaccine for long-term protection							
CC against plague. Expression by gut-colonising bacterial transformants							
CC Sequence 329 AA;							

Query Match	61.6%	Score 1641.5	DB 1	Length 329
Best Local Similarity	99.4%	Pred. NO.1.2e-110		
Matches 326	Conservative 0	Mismatches 1	Indels 1	Gaps 1
OY	194	EFMIRAYEONPHQFTEDLEKVRVEOLTGHGSSVLEELVOLVYDKNIDISIKYDPRKDSEY	253	

Query Match	61.6%	Score 1641	DB 1	Length 329
Best Local Similarity	99.7%	Pred. No. 1.3e-110		
Matches 324; Conservative	0	Mismatches 1	Indels 0	Gaps 0

DB	QY	DB	QY
5	1RAYEQNPQHFIEDLEKRVKEQLTGSSSVLEELVOLVKDKNDISRYDPRKQSEFPAN	64	
257	RYITDDIELKLIILFLPEDIILGGHYDNLONGIRKRVKEFLSSPNTOMELAFMAV	316	
65	RYITDDIELKLIILFLPEDIILGGHYDNLONGIRKRVKEFLSSPNTOMELAFMAV	124	

OY 317 MHFSLTADRIDDDILKIVYDSNMHHGDARSKRLRELAETALAKIYSYQAENKHLSS 376
|
DB 125 MHFSLTADRIDDDILKIVYDSNMHHGDARSKRLRELAETALAKIYSYQAENKHLSS 184
OY 377 GTINIHDSINIMDMKNLYGYTDEEIFKASAEKYLEKMPQTTIOYDGSSEKIKVSIKDFLG 436
|
DB 185 GTINIHDSINIMDMKNLYGYTDEEIFKASAEKYLEKMPQTTIOYDGSSEKIKVSIKDFLG 244
OY 437 SENKRTGALGNKNSYKYNKNNELSHFATTCSDKSRLPNDLVSKOTQLOSDITSRPNFA 496
|
DB 245 SENKRTGALGNKNSYKYNKNNELSHFATTCSDKSRLPNDLVSKOTQLOSDITSRPNFA 304
OY 497 IEALNRFIOKDYDVQMRLLDDTSGR 521
|
DB 305 IEALNRFIOKDYDVQMRLLDDTSGR 329

RESULT 6
R79962
ID R79962 standard; Protein: 329 AA.
AC R79962;
DT 18-APR-1996 (first entry)
DE Partial lcrv (V antigen) of Y. pestis.
KW lcrv; V antigen; virulence; plague; vaccine; epitope.
OS Yersinia pestis.
PN W09524475-A1.
PD 14-SEP-1995;
PF 06-MAR-1995; G00481.
PR 08-MAR-1994; GB-004577.
PA (MINA) UK SEC FOR DEFENCE.
PI Leary SEC, Tiltball RW, Williamson ED, Leary SE;
DR WPI: 95-328268/42.
N-PSDB: T04223.
FT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PT oral or parenteral vaccines for protection against plague
PS Claim 6; Page 15-16; 25PP; English.
CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrv) encoding
CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrv) encoding
CC all or a protective epitopic part of the mature V protein of Yersinia
CC pests. The protein was expressed as a fusion protein with maltose
CC binding protein or glutathione-S-transferase in 3 different plasmid
CC vectors. Y. pestis is the highly virulent causative organism of plague
CC in a wide range of animals, including man. The V antigen (lcrv) is an
CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
CC The V antigen is postulated to act as a virulence antigen, and
CC transformed microorganisms contg. recombinant DNA encoding a V antigen
CC protein/peptide are useful in vaccines to protect against plague.
CC Sequence 329 AA;
SQ

Query Match 61.4%; Score 1638; DB 1; Length 329;
Best Local Similarity 99.4%; Pred. No. 2.1e-110;
Matches 323; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 197 IRAYEONPHQFIEDLEKRVVEQLTGHSSVLEELVQLVKDKNIDISIKYDPKDESEFAN 256
|
DB 5 IRAYEONPHQFIEDLEKRVVEQLTGHSSVLEELVQLVKDKNIDISIKYDPKDESEFAN 64
OY 257 RYITDDIELKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLESSPNTONELRAFMAY 316
|
DB 65 RYITDDIELKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLESSPNTONELRAFMAY 124
OY 317 MHFSLTADRIDDDILKIVYDSNMHHGDARSKRLRELAETALAKIYSYQAENKHLSS 376
|
DB 125 MHFSLTADRIDDDILKIVYDSNMHHGDARSKRLRELAETALAKIYSYQAENKHLSS 184
OY 377 GTINIHDSINIMDMKNLYGYTDEEIFKASAEKYLEKMPQTTIOYDGSSEKIKVSIKDFLG 436
|
DB 185 GTINIHDSINIMDMKNLYGYTDEEIFKASAEKYLEKMPQTTIOYDGSSEKIKVSIKDFLG 244
OY 437 SENKRTGALGNKNSYKYNKNNELSHFATTCSDKSRLPNDLVSKOTQLOSDITSRPNFA 496
|
DB 245 SENKRTGALGNKNSYKYNKNNELSHFATTCSDKSRLPNDLVSKOTQLOSDITSRPNFA 304

OY 497 IEALNRFIOKDYDVQMRLLDDTSGR 521
|
DB 305 IEALNRFIOKDYDVQMRLLDDTSGR 329

RESULT 7
R76528
ID R76528 standard; Protein: 170 AA.
AC R76528;
DT 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen.
KW vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
OS Yersinia pestis.
PN W09518231-A1.
PD 06-JUL-1995;
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA) UK SEC FOR DEFENCE.
PI Howells A, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
DR WPI: 95-246396/32.
N-PSDB: Q92819.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.
PS Disclosure; Page 20; 2/PP; English.
CC The sequence represents the Y. pestis cafi (F1) antigen expressed
CC from plasmid pFOR1b. The DNA construct can be used to transform
CC human or animal gut colonizing microorganisms, specifically
CC attenuated Salmonella typhimurium or Salmonella typhi. The
CC transformed microorganisms can be used as live/attenuated vaccines
CC which induce immune responses at mucosal surfaces. The vaccines
CC provide protection against infection with Y. pestis, and are
CC parenterally and orally active vaccines offering protection
CC against bubonic and pneumonic plague.
CC Sequence 170 AA;
SQ

Query Match 31.9%; Score 850; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 3.6e-54;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 24 MKKISSVIAIALFGIATANAADLTATATATLVEPRKITLTKREGAPITMDNGNIDT 83
|
DB 1 MKKISSVIAIALFGIATANAADLTATATATLVEPRKITLTKREGAPITMDNGNIDT 60
OY 84 ELVGTILFLGGYKGTGTSTSVNETDAADPMYLTFTSODGNNHOFPTTVIGKDSRDPIS 143
|
DB 61 ELVGTILFLGGYKGTGTSTSVNETDAADPMYLTFTSODGNNHOFPTTVIGKDSRDPIS 120
OY 144 PRVNGENLVGDVYLVATGSDGFVRSISGSKGLAAGKYTDAVTYVSNO 193
|
DB 121 PRVNGENLVGDVYLVATGSDGFVRSISGSKGLAAGKYTDAVTYVSNO 170

RESULT 8
W01043
ID W01043 standard; Protein: 170 AA.
AC W01043;
DT 28-DEC-1996 (first entry)
DE Y. pestis F1 antigen (including signal peptide).
KW plague; vaccine; genetic immunisation; F1 antigen; cafi;
OS Yersinia pestis strain GB.
FH Key location/Qualifiers
FT peptide 1..21
PN W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1995; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.

PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tildall RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR N-PSDB: T38248.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 epitopic parts - useful in vaccine for protection against plague
 PS Example 2; Page 61-62; 98pp; English.
 CC The F1 antigen (W01043), including the signal peptide, of Yersinia
 CC pestis was produced from a DNA sequence (T38248) obtd. by PCR
 CC amplification (see also T38257-58) of Y. pestis DNA. Expression
 CC of the F1 antigen (see also W01042) by gut-colonising organisms in
 CC the form of live vaccines can be used to protect an animal,
 CC including humans, against plague.
 SQ Sequence 170 AA;

Query Match 31.9%; Score 850; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 3.6e-54;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAIAALFGTIANANADLASTATATLVEPARITLTYEGAPITIMNGNIDT 83
 DB 1 MKKISSVIAIAALFGTIANANADLASTATATLVEPARITLTYEGAPITIMNGNIDT 60
 QY 84 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 143
 DB 61 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 120
 QY 144 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 193
 DB 121 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 170

RESULT 9
 W59782
 ID W59782 standard; Protein; 170 AA.
 AC W59782;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYpFlsec170.
 KW F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 DR N-PSDB: V41594.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 52; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 170 AA;

Query Match 31.9%; Score 850; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 3.6e-54;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAIAALFGTIANANADLASTATATLVEPARITLTYEGAPITIMNGNIDT 83
 DB 1 MKKISSVIAIAALFGTIANANADLASTATATLVEPARITLTYEGAPITIMNGNIDT 60
 QY 84 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 143
 DB 61 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 120
 QY 144 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 193

DB 121 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 170

RESULT 10
 W59783
 ID W59783 standard; Protein; 170 AA.
 AC W59783;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen nYpFlsec510.
 KW F1 antigen; plasmid; vaccine; plague; ds.
 OS Yersinia pestis.
 FH Key Location/Qualifiers
 FT CDS 17..532
 FT /tag- a
 FT /product- "F1 antigen"
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 DR N-PSDB: V41596.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Disclosure; Pages 53-54; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 170 AA;

Query Match 31.9%; Score 850; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 3.6e-54;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAIAALFGTIANANADLASTATATLVEPARITLTYEGAPITIMNGNIDT 83
 DB 1 MKKISSVIAIAALFGTIANANADLASTATATLVEPARITLTYEGAPITIMNGNIDT 60
 QY 84 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 143
 DB 61 ELVGTILGKYGKGTSTSVNFTDAAGDPMYLFTSODGNNHOFTRKVGKDSRDPDIS 120
 QY 144 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 193
 DB 121 PKVGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKTYDAVTVTSNQ 170

RESULT 11
 W59788
 ID W59788 standard; Protein; 149 AA.
 AC W59788;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen pYpFlmat149.
 KW F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 DR N-PSDB: V41609.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 63; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to

CC Produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 149 AA;

Query Match 28.3%; Score 755; DB 1; Length 149;
 Best Local Similarity 100.0%; Pred. No. 2e-47;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 ADLTAATATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSY 104
 DB 1 ADLTAATATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSY 60
 OY 105 NETDAAGDPMYLTFTSQDGNHOFITTKVIGKDSRDPDISPKVNGENLVGDDVVLATGSD 164
 DB 61 NETDAAGDPMYLTFTSQDGNHOFITTKVIGKDSRDPDISPKVNGENLVGDDVVLATGSD 120
 OY 165 FFVRISIGSKGKLAAGKYTDAVTVVSNQ 193
 DB 121 FFVRISIGSKGKLAAGKYTDAVTVVSNQ 149

RESULT 12
 W59787

ID W59787 standard; Protein: 150 AA.

AC W59787.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYPFmat150.
 KW F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HEK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-333331/29.
 DR N-PSDB: V41600.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 60; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 150 AA;

Query Match 28.3%; Score 755; DB 1; Length 150;
 Best Local Similarity 100.0%; Pred. No. 2e-47;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 ADLTAATATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSY 104
 DB 2 ADLTAATATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSY 61
 OY 105 NETDAAGDPMYLTFTSQDGNHOFITTKVIGKDSRDPDISPKVNGENLVGDDVVLATGSD 164
 DB 62 NETDAAGDPMYLTFTSQDGNHOFITTKVIGKDSRDPDISPKVNGENLVGDDVVLATGSD 121
 OY 165 FFVRISIGSKGKLAAGKYTDAVTVVSNQ 193
 DB 122 FFVRISIGSKGKLAAGKYTDAVTVVSNQ 150

RESULT 13
 R76526
 ID R76526 standard; Protein: 151 AA.
 AC R76526;
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis cafi (F1) antigen.

KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 KW bubonic plague; pneumonic plague.
 OS Yersinia pestis.
 PN W09518231-A1.
 PD 06-JUL-1995.
 PF 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
 DR WPI: 95-246396/32.
 DR N-PSDB: G92817.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure; Page 16; 27pp; English.
 CC The sequence represents the Y. pestis cafi (F1) antigen expressed
 CC from plasmid pFGAL2a. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated Salmonella typhimurium or Salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 151 AA;

Query Match 28.3%; Score 755; DB 1; Length 151;
 Best Local Similarity 98.0%; Pred. No. 2e-47;
 Matches 148; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 43 NADITATATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSY 102
 DB 1 SSADITATATATLVEPARTITLYKKGAPITIMDNGNIDELVGLTLLGKYKGTSTSY 60
 OY 103 SYNFTDAAGDPMYLTFTSQDGNHOFITTKVIGKDSRDPDISPKVNGENLVGDDVVLATGS 162
 DB 61 SYNFTDAAGDPMYLTFTSQDGNHOFITTKVIGKDSRDPDISPKVNGENLVGDDVVLATGS 120
 OY 163 ODFFVRISIGSKGKLAAGKYTDAVTVVSNQ 193
 DB 121 ODFFVRISIGSKGKLAAGKYTDAVTVVSNQ 151

RESULT 14
 W01042

ID W01042 standard; Protein: 151 AA.

AC W01042.
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1 antigen.
 KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;
 KW V antigen.
 OS Yersinia pestis strain GB.
 PN W09628551-A1.
 PD 19-SEP-1996.
 PF 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MTNA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR N-PSDB: T38244.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Disclosure; Page 43-45; 98pp; English.
 CC Yersinia pestis F1 antigen (W01042) is capable of evoking protective
 CC immune responses in animals. A combined vaccine of F1 and V
 CC antigens (see also W01040-41) can at least match the protection
 CC afforded by live attenuated EV76 vaccine without any of the hazards
 CC that have kept the EV vaccine from general use. The F1 antigen is
 CC produced using the cafi gene (see also T38244) obtd. from Y.
 CC pestis by PCR amplification. It can also be prepd. as a fusion with

CC V antigen (see also W01044-45) and expressed by attenuated AroA or
 CC C Salmonella typhi as a live vaccine for long-term protection
 CC against plague. Expression by gut-colonising bacterial transformants
 CC produces a protective response against Y. pestis.
 SQ Sequence 151 AA;

Search completed: August 22, 2000, 16:46:38
 Job time: 2510 sec

Query Match 28.3%; Score 755; DB 1; Length 151;
 Best Local Similarity 98.0%; Pred. No. 2e-47;
 Matches 148; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 43 NAAULTASTATATATVEPARITITKEGAPITIMDNGNIDTELLVGTLLGGYKGTST 102
 DB 1 SSADLTASTATATATVEPARITITKEGAPITIMDNGNIDTELLVGTLLGGYKGTST 60
 QY 103 SVNFTDAGDPMYLTFTSODGNHOFITKVGKDSRDPDISPKVNGENLVGDVYLATGS 162
 DB 61 SVNFTDAGDPMYLTFTSODGNHOFITKVGKDSRDPDISPKVNGENLVGDVYLATGS 120
 QY 163 QDFEVRISIGSKGKLAGKYTDATVTVSNQ 193
 DB 121 QDFEVRISIGSKGKLAGKYTDATVTVSNQ 151

RESULT 15

R76527
 ID R76527 standard; Protein; 151 AA.
 AC R76527;
 DE 17-DEC-1995 (first entry)
 KW Yersinia pestis cafi (F1) antigen.
 DE Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 KW bubonic plague; pneumonic plague.
 OS Yersinia pestis.
 PN W09518231-A1.
 PD 06-JUL-1985.
 PF 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A. Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
 DR WPI: 95-246396/32.
 DR N-PSDB: Q92818.
 PT DNA constructs capable of transforming microorganisms - which can be
 used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure: Page 18; 27pp; English.
 CC The sequence represents the Y. pestis cafi (F1) antigen expressed
 CC from plasmid pF51G3a. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated Salmonella typhimurium or Salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 151 AA;

Query Match 28.2%; Score 753; DB 1; Length 151;
 Best Local Similarity 99.3%; Pred. No. 2.8e-47;
 Matches 148; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 45 ADLTASTATATATVEPARITITKEGAPITIMDNGNIDTELLVGTLLGGYKGTSTSV 104
 DB 3 ADLTASTATATATVEPARITITKEGAPITIMDNGNIDTELLVGTLLGGYKGTSTSV 62
 QY 105 NETDAAGDPMYLTFTSODGNHOFITKVGKDSRDPDISPKVNGENLVGDVYLATGSOD 164
 DB 63 NETDAAGDPMYLTFTSODGNHOFITKVGKDSRDPDISPKVNGENLVGDVYLATGSOD 122
 QY 165 FFEVRISIGSKGKLAGKYTDATVTVSNQ 193
 DB 123 FFEVRISIGSKGKLAGKYTDATVTVSNQ 151

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 16:29:03 ; Search time 22.92 seconds

(without alignments)
348.451 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 2666
1 MGHNNHHHHHHSSGHIDD.....RFIOKYDSVWQRLLDPTSGK 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 152396 seqs, 15329161 residues

Total number of hits satisfying chosen parameters: 152396

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2480.5	93.0	501	2	US-08-913-477-23
2	2378.5	89.2	480	2	US-08-913-477-17
3	1641.5	61.6	329	2	US-08-913-477-2
4	1641	61.6	329	2	US-08-913-477-4
5	844	31.7	170	2	US-08-913-477-21
6	755	28.3	151	2	US-08-913-477-11
7	159	6.0	24	3	US-08-780-496-8
8	144	5.4	462	1	US-08-480-604A-26
9	144	5.4	462	2	US-08-405-496A-26
10	132	5.0	53	2	US-08-651-818A-19
11	132	5.0	54	2	US-08-651-818A-23
12	125.5	4.7	799	1	US-08-188-228-42
13	125.5	4.7	799	1	US-08-332-638-42
14	124	4.7	1683	3	US-08-735-587-163
15	124	4.6	23	1	US-08-480-604A-24
16	122	4.6	23	1	US-08-405-496A-24
17	122	4.6	323	1	US-08-591-989-5
18	119.5	4.5	793	1	US-08-188-228-54
19	119.5	4.5	793	1	US-08-332-643-48
20	119.5	4.5	793	1	US-08-332-638-54
21	119	4.5	21	2	US-08-651-818A-21
22	118.5	4.4	532	1	US-08-188-228-44
23	118.5	4.4	532	1	US-08-332-638-44
24	118	4.4	800	3	US-08-776-265-3
25	115	4.3	859	1	US-08-053-614-2
26	115	4.3	859	1	US-08-316-397B-2
27	115	4.3	859	2	US-09-034-306-2
28	115	4.3	859	4	PCT-US93-09782-2

29	115	4.3	1181	1	US-08-053-614-4	Sequence 4, Appl1
30	115	4.3	1181	1	US-08-316-397B-4	Sequence 4, Appl1
31	115	4.3	1181	2	US-09-034-306-4	Sequence 4, Appl1
32	115	4.3	1181	4	PCT-US93-09782-4	Sequence 4, Appl1
33	113.5	4.3	792	3	US-08-433-522A-8	Sequence 8, Appl1
34	113	4.2	1354	3	US-08-685-871-2	Sequence 2, Appl1
35	112	4.2	711	3	US-08-949-588-2	Sequence 2, Appl1
36	112	4.2	2329	4	US-08-755-587-16	Sequence 16, Appl1
37	112	4.2	2938	4	PCT-US94-00198-3	Sequence 3, Appl1
38	112	4.2	3111	2	US-08-460-309-4	Sequence 4, Appl1
39	112	4.2	3111	2	US-08-125-077-4	Sequence 4, Appl1
40	112	4.2	3418	2	US-08-639-501-2	Sequence 2, Appl1
41	112	4.2	3418	3	US-09-044-946-2	Sequence 2, Appl1
42	111.5	4.2	736	1	US-07-688-352C-24	Sequence 24, Appl1
43	111.5	4.2	736	2	US-08-474-379C-24	Sequence 24, Appl1
44	111.5	4.2	736	3	US-09-146-249A-24	Sequence 24, Appl1
45	111.5	4.2	736	4	PCT-US91-02714-23	Sequence 23, Appl1

ALIGNMENTS

RESULT 1
US-08-913-477-23
; Sequence 23, Application US/08913477
; Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tiltball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Gable Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-23

Query Match 93.0%; Score 2480.5; DB 2; Length 501;
Best local Similarity 98.8%; Pred. No. 3,7e-187;
Matches 495; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 24 MKRISSVIALFGLTATANAADLSTATAFLVEPARITLTKKGAPITINDNGNIDT 83
DB 1 MKRISSVIALFGLTATANAADLSTATAFLVEPARITLTKKGAPITINDNGNIDT 60
QY 84 ELIVGLTLGGYKGTGTSVNTDAAGDPMTLFTSODSNNHQTFTKVGKSRPDIS 143
DB 61 ELIVGLTLGGYKGTGTSVNTDAAGDPMTLFTSODSNNHQTFTKVGKSRPDIS 120
QY 144 PKYNGENLVGDDVYLATGSDFFVRSIGSGKGLAGKYDVAVTYSNO---EFMIRAY 200
DB 121 PKYNGENLVGDDVYLATGSDFFVRSIGSGKGLAGKYDVAVTYSNOGSIKGRIRAY 180
QY 201 EONPQHIEDLEKRVYEQLTGHSSVLEELVQVVKDNIDISIKYDPRKDEVFANRVT 260
DB 181 EONPQHIEDLEKRVYEQLTGHSSVLEELVQVVKDNIDISIKYDPRKDEVFANRVT 240
QY 261 DDEILKTLIAYFLPDTILKGGHYDNOLONGIKRYKEFLESSPNTOMELRAVMVHFS 320
DB 241 DDEILKTLIAYFLPDTILKGGHYDNOLONGIKRYKEFLESSPNTOMELRAVMVHFS 300
QY 321 LPADRIDDDILKIVDSMNHGDAKRLRELAELTAELKIVSIOAEIKRHLSSGTTIN 380
DB 301 LPADRIDDDILKIVDSMNHGDAKRLRELAELTAELKIVSIOAEIKRHLSSGTTIN 360
QY 381 IHDKSNLMDKNLYGTDEEIFKASAEYKILEKMPOTTIQVDSEKKIYSIKDFLSENK 440
DB 361 IHDKSNLMDKNLYGTDEEIFKASAEYKILEKMPOTTIQVDSEKKIYSIKDFLSENK 420
QY 441 RIGALNLKNSYSYNKDNHLSHFAITCSDKSRPLNDVYSQKTTQSDITSRNSAIEAL 500
DB 421 RIGALNLKNSYSYNKDNHLSHFAITCSDKSRPLNDVYSQKTTQSDITSRNSAIEAL 480
QY 501 NREIQYDSVMORLDDTSGK 521
DB 481 NREIQYDSVMORLDDTSGK 501

RESULT 2

US-08-913-477-17
Sequence 17, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-17

Query Match 89.2%; Score 2378.5; DB 2; Length 480;
Best local Similarity 98.5%; Pred. No. 3.6e-179;
Matches 473; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 45 ADLSTATAFLVEPARITLTKKGAPITINDNGNIDTELLVGLTLGGYKGTGTSV 104
DB 1 ADLSTATAFLVEPARITLTKKGAPITINDNGNIDTELLVGLTLGGYKGTGTSV 60
QY 105 NPTDAAGDPMTLFTSODSNNHQTFTKVGKSRPDISPKYNGENLVGDDVYLATGSD 164
DB 61 NPTDAAGDPMTLFTSODSNNHQTFTKVGKSRPDISPKYNGENLVGDDVYLATGSD 120
QY 165 FVRSIGSGKGLAGKYDVAVTYSNO---EFMIRAYEONPQHIEDLEKRVYEQLTG 221
DB 121 FVRSIGSGKGLAGKYDVAVTYSNOGSIKGRIRAYEONPQHIEDLEKRVYEQLTG 180
QY 222 HGSVLEELVQVVKDNIDISIKYDPRKDEVFANRVTDDIELKTLIAYFLPDTILK 281
DB 181 HGSVLEELVQVVKDNIDISIKYDPRKDEVFANRVTDDIELKTLIAYFLPDTILK 240
QY 282 GGHYDNOLONGIKRYKEFLESSPNTOMELRAVMVHFSLPADRIDDDILKIVDSMNH 341
DB 241 GGHYDNOLONGIKRYKEFLESSPNTOMELRAVMVHFSLPADRIDDDILKIVDSMNH 300
QY 342 GDAKRLRELAELTAELKIVSIOAEIKRHLSSGTTINHKSNLMDKNLYGTDEEI 401
DB 301 GDAKRLRELAELTAELKIVSIOAEIKRHLSSGTTINHKSNLMDKNLYGTDEEI 360
QY 402 FRASAEYKILEKMPOTTIQVDSEKKIYSIKDFLSENKRTGALNLKNSYSYNKDNEL 461
DB 361 FRASAEYKILEKMPOTTIQVDSEKKIYSIKDFLSENKRTGALNLKNSYSYNKDNEL 420
QY 462 SHFATTCSDKSRPLNDVYSQKTTQSDITSRNSAIEALNREIQYDSVMORLDDTSGK 521
DB 421 SHFATTCSDKSRPLNDVYSQKTTQSDITSRNSAIEALNREIQYDSVMORLDDTSGK 480

RESULT 3

US-08-913-477-2
Sequence 2, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.

APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-2

Query Match 61.6%; Score 1641.5; DB 2; Length 329;
Best Local Similarity 99.4%; Pred. No. 1.6e-121;
Matches 326; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 194 EFMIRAYQNPOHFIEDLEKRVVEQLTGHSSVLEELVQLVKDKNIDISIKYPRKDESEV 253
Db 3 EF-IRAYEQNPOHFIEDLEKRVVEQLTGHSSVLEELVQLVKDKNIDISIKYPRKDESEV 61
QY 254 FANRVTITDIELKKILAYFLPEDTILKGHYDNOLONGIKRVKFELESSPNTOWELRAF 313
Db 62 FANRVTITDIELKKILAYFLPEDTILKGHYDNOLONGIKRVKFELESSPNTOWELRAF 121
QY 314 MAVMFSIADRIDDDILKVIYDSNNHHGDARSKLREELAEITAEIKTISVQAEINKHL 373
Db 122 MAVMFSIADRIDDDILKVIYDSNNHHGDARSKLREELAEITAEIKTISVQAEINKHL 181
QY 374 SSSGTINIHDKSINIMDKNLVGTDEEITFKASAEYKILKKMPOTTIOVDGSEKKIYSIKD 433
Db 182 SSSGTINIHDKSINIMDKNLVGTDEEITFKASAEYKILKKMPOTTIOVDGSEKKIYSIKD 241
QY 434 FLGSNKRRTGALGNLKNSTYKNNKNNELSHFATITCSDKSRPLNDLVSOQTTLDSITSRF 493
Db 242 FLGSNKRRTGALGNLKNSTYKNNKNNELSHFATITCSDKSRPLNDLVSOQTTLDSITSRF 301

QY 494 NSAIKALNRFIOKYDSVWQRLDDTSGK 521
Db 302 NSAIKALNRFIOKYDSVWQRLDDTSGK 329
RESULT 4
US-08-913-477-4
Sequence 4, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-4
Query Match 61.6%; Score 1641; DB 2; Length 329;
Best Local Similarity 99.7%; Pred. No. 1.8e-121;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 197 IRAYEQNPOHFIEDLEKRVVEQLTGHSSVLEELVQLVKDKNIDISIKYPRKDESEV 256
Db 5 IRAYEQNPOHFIEDLEKRVVEQLTGHSSVLEELVQLVKDKNIDISIKYPRKDESEV 64
QY 257 RVTITDIELKKILAYFLPEDTILKGHYDNOLONGIKRVKFELESSPNTOWELRAF 316
Db 65 RVTITDIELKKILAYFLPEDTILKGHYDNOLONGIKRVKFELESSPNTOWELRAF 124

QY 317 MRSITADRIDDDILKVIYDSNHHGDAKSKLEELAEELAEIKIYSVIOAEINKHLSSS 376
|
Db 125 MHSLSIADRIDDDILKVIYDSNHHGDAKSKLEELAEELAEIKIYSVIOAEINKHLSSS 184
QY 377 GTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVGSEKKIYSINDFLG 436
|
Db 185 GTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVGSEKKIYSINDFLG 244
QY 437 SENKRGALGNLKNYSYNKNDELPHATGSKSRPLNDLYSOKTQLOSDITSRFSNSA 496
|
Db 245 SENKRGALGNLKNYSYNKNDELPHATGSKSRPLNDLYSOKTQLOSDITSRFSNSA 304
QY 497 IEALNRFIOKYDSVMORLLDDTSGK 521
|
Db 305 IEALNRFIOKYDSVMORLLDDTSGK 329

RESULT 5

US-08-913-477-21
; Sequence 21, Application US/08913477
; Patent No. 5985285
; GENERAL INFORMATION:
; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Leary, Sophie E.C.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Bennett, Alice M.
; TITLE OF INVENTION: VACCINES FOR PLAGUE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,477
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 424
; APPLICATION NUMBER: PCT/GB96/00571
; FILING DATE: 13-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9505059.7
; FILING DATE: 13-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9518946.0
; FILING DATE: 15-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9524825.8
; FILING DATE: 05-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 124-599
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4100
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-913-477-21

Query Match 31.7%; Score 844; DB 2; Length 170;
Best Local Similarity 99.4%; Pred. No. 2.9e-59;
Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 MKRISVIAIAFGTATNADLTASTATATVTEPARITLTYEGAPITMDNGNDIT 83
|
Db 1 MKRISVIAIAFGTATNADLTASTATATVTEPARITLTYEGAPITMDNGNDIT 60
QY 84 ELVGLTLGKGTGTSTSVNFTDAAGDPMYLFTTSODGNHOFITVYIGKDSRDFDIS 143
|
Db 61 ELVGLTLGKGTGTSTSVNFTDAAGDPMYLFTTSODGNHOFITVYIGKDSRDFDIS 120
QY 144 PVRNGENLVGDVYVATGSGDFVRSIGSGKGLAGKYTDAVTVTSNQ 193
|
Db 121 PVRNGENLVGDVYVATGSGDFVRSIGSGKGLAGKYTDAVTVTSNQ 170

RESULT 6

US-08-913-477-11
; Sequence 11, Application US/08913477
; Patent No. 5985285
; GENERAL INFORMATION:
; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Leary, Sophie E.C.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Bennett, Alice M.
; TITLE OF INVENTION: VACCINES FOR PLAGUE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,477
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 424
; APPLICATION NUMBER: PCT/GB96/00571
; FILING DATE: 13-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9505059.7
; FILING DATE: 13-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9518946.0
; FILING DATE: 15-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9524825.8
; FILING DATE: 05-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 124-599
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4100
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-913-477-11

Query Match	28.3%	Score 755;	DB 2;	Length 151;
Best Local Similarity	98.0%;	Pred. No. 2.4e-52;		
Matches 148;	Conservative 3;	Mismatches 0;	Indels 0;	Gaps 0;

Oy 43 MADDLTAATAATVEPARTITLYKSGAPITINDNGNIDTELVGTLGGYKGTGTST 102
 Db 1 SSADLTASTATATATVEPARTITLYKSGAPITINDNGNIDTELVGTLGGYKGTGTST 60

QY 103 SVAFTDAAGDPMYLTFTSQDGNHHQFTTKVIGKDSRDFDISPKVNGENLVGDDVYLATGS 1622
|||||
D6 61 SVAFTDAAGDPMYLTFTSQDGNHHQFTTKVIGKDSRDFDISPKVNGENLVGDDVYLATGS 1200

```
Qy 163 QDFVRSIGSKGKLAAGKYTDAVTVTVSNO 1933
    |||||
Db 121 QDFVRSIGSKGKLAAGKYTDAVTVTVSNO 1511
```

RESULT 7
US-08-780-496-8
Sequence 8, Application US/08780496

APPLICANT: Avishashnazi, Aram Chuntcharepat, Kyung Jin Kim
 TITLE OF INVENTION: Apo-2 Ligand
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinpatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/780,496
FILING DATE: 08-Jan-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

```

? REGISTRATION NUMBER: 35,600
? REFERENCE/DOCKET NUMBER: P0978P1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415/223-5416
? TELEFAX: 415/952-9881
? TELER: 910/371-7168
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 24 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
US-08-780-496-8
```

Query Match	6.0%	Score 159;	DB 3;	Length 24;
Best Local Similarity	100.0%	Pred. No. 1.1e-06;		
Matches 24; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

```
QY 1 MGNHHHHHHHSSGHIDDDDKHM 24
    |||||||
Db 1 MGNHHHHHHHSSGHIDDDDKHM 24
```

RESULT 8
US-08-480-604A-26
; Sequence 26, Application US/08480604A
; Patent No. 5736139
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.

APPLICANT: PADHAYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.3C
CURRENT APPLICATION DATA: 000000 0000

APPLICATION NUMBER: US/08/480,604
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA.

APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763

```

TELECOMMUNICATION INFORMATION
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
Type: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
OS-08-480-604A-26

```

Query Match	5.4%;	Score 144;	DB 1;	Length 462;
Best Local Similarity	17.8%;	Pred. No. 0.0012;		
Matches	86;	Conservative	89;	Mismatches 139; Indels 168; Gaps 22;

```
QY      1 MGNHHHHHHHSSGHIIDDDDKKKKSSVIALALETITANNAADITASTTATATLVER 60
        |||||            |:::|          |:-|          :::
        |||||            |:::|          |:-|          :::
        |||||            |:::|          |:-|          :::
Db      1 MGNHHHHHHHSSGHI - EGRMAASARL-----LSTETETIKNIINT 42
```

```

QY      61 ARITLTKEGAPITMDNGNIDELLVGTITLGKYKTGTSTSVNFDAAGDPWLTFTS 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      43 SIMLARE-----SNHLIDLSRYASKINIG-----SKYNF-----DPI----- 75
      Db

```

QY 121 QDGNHNETTKRVGKSDREDPISPKNGEYKLVGGDYVLANGSDPEFRSISGSGKRLAAG 180
76 -DKNQIQ-----LFWLESKEIVLIK-----NAIYNSKTEFNSTSFWR-----IP 116
Db 181 KYTDAVTATVSNQEFMIRATEQNPO-----HFIEDLKRVRREQLTGSGSVLEE 229

DB 117 KYFN--SISLNNEYTIINCENNSGKVSUNYGEIITWLTQDQEI-K-----QRVYFK 166
QY 230 LVOLVK-----DKNIDISIKYDPRKDEVFANRVITDDIELLKLILAYFLPEDTILKGGH 284
DB 167 YSQMINISDYINRMIFVTITNNRLNLSKIYINGRLIDQ-----KPI-----207
QY 285 YDNQOLNGIKRVKFEFLSSPTQWEL-----RAFVAVHFSLTADRIDDLKVIYDSM 338
DB 208 --SNLGN-----IHASNNIMEKLDGCRDTHRYIWIKEFNLEFDEKLEKEIKDLQ 257
QY 339 NHGDARSKRLRELAETALKEIYSVIOAEINKHLSSSGTINIHD--KSLNLDKKNLYG 396
DB 258 SNGS-----ILKDFWGDLYQDKPIYM-----LNLDPKRYVDVNNNGIRGY 299
QY 397 TDEIFKASAEYKILEKMPQTTIOVDGSEKKIYSINDFJGSENNKRTGALGNLKNYSYK 456
DB 300 -----MYLKGPRGSVMTNTIYLNSSLYR-GTKFLIKKVASGNK 336
QY 457 DN 458
DB 337 DN 338

RESULT 9
US-08-405-496A-26
Sequence 26, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
TITLE OF INVENTION: NEUROTOXIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-405-496A-26
Query Match 5.4%; Score 144; DB 2; Length 462;
Best Local Similarity 17.8%; Pred. No. 0.0012;
Matches 86; Conservative 89; Mismatches 139; Indels 168; Gaps 22;
QY 1 MGHHHHHHHHHSSGHIDDDKMKKISSVIALALGTATANAADLTASTATATATVLP 60
DB 1 MGHHHHHHHHHSSGHI--EGRMASMARL-----LSFTTEIKINIWT 42
QY 61 ARTITLYKEGAPITTDNGNDIDELLVGLTLAGYGTGTTTSVNTDAGDPMTLTFTS 120
DB 43 SLNRYE-----SNHLIDLSRYASKINIG--SKVNF-----DPI-----75
QY 121 QDNHNOFTTKVIGKSDRFDISPKVNGENLVDVYLATGSDPEFVRISGSGKTLANG 180
DB 76 -DKNQIQ-----LNFESSKIEYIK--NAIYNSMYENFSTFWIR-----IP 116
QY 181 KYTDAVTYVSNQEFMIRAYEONPQ-----HFTEDLEKRVEDLTGHSSVLEE 229
DB 117 KYFN--SISLNNEYTIINCENNSGKVSUNYGEIITWLTQDQEI-K-----QRVYFK 166
QY 230 LVOLVK-----DKNIDISIKYDPRKDEVFANRVITDDIELLKLILAYFLPEDTILKGGH 284
DB 167 YSQMINISDYINRMIFVTITNNRLNLSKIYINGRLIDQ-----KPI-----207
QY 285 YDNQOLNGIKRVKFEFLSSPTQWEL-----RAFVAVHFSLTADRIDDLKVIYDSM 338
DB 208 --SNLGN-----IHASNNIMEKLDGCRDTHRYIWIKEFNLEFDEKLEKEIKDLQ 257
QY 339 NHGDARSKRLRELAETALKEIYSVIOAEINKHLSSSGTINIHD--KSLNLDKKNLYG 396
DB 258 SNGS-----ILKDFWGDLYQDKPIYM-----LNLDPKRYVDVNNNGIRGY 299
QY 397 TDEIFKASAEYKILEKMPQTTIOVDGSEKKIYSINDFJGSENNKRTGALGNLKNYSYK 456
DB 300 -----MYLKGPRGSVMTNTIYLNSSLYR-GTKFLIKKVASGNK 336
QY 457 DN 458
DB 337 DN 338

RESULT 10
US-08-651-818A-19
Sequence 19, Application US/08651818A
Patent No. 5948889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
TITLE OF INVENTION: ANTIMICROBIALS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroli, Peter G.

REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-19

Query Match 5.0%; Score 132; DB 2; Length 53;
Best Local Similarity 73.3%; Pred. No. 0.00045;
Matches 22; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MGHNNHHHHSSGHI-----DDDDK 22
DB 1 MGHNNHHHHSSGHI-----DDDDK 30

RESULT 11
US-08-651-818A-23

Sequence 23, Application US/08651818A
Patent No. 5948889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
APPLICANT: Hale, Cynthia A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-23

Query Match 5.0%; Score 132; DB 2; Length 54;
Best Local Similarity 73.3%; Pred. No. 0.00047;
Matches 22; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 MGHNNHHHHSSGHI-----DDDDK 22
DB 1 MGHNNHHHHSSGHI-----DDDDK 22

DB 1 MGHNNHHHHSSGHI-----DDDDK 30

RESULT 12

US-08-188-228-42
Sequence 42, Application US/08188228
Patent No. 5597725
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,228
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/049,460
FILING DATE: 19 APR 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,643
FILING DATE: 17 APR 1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5597725and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31340
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-188-228-42

Query Match 4.7%; Score 125.5; DB 1; Length 799;
Best Local Similarity 21.3%; Pred. No. 0.078;
Matches 107; Conservative 80; Mismatches 199; Indels 117; Gaps 25;

QY 19 DDDKMKKISSVIALALGTIATAAADLTASTATL--VEPARILT-----YKGA 71
DB 88 DDDPSKRIKIKYLSGDSAGTIFQIN--DITGDIHAKIRLBREKREYLLTAQAWDMENK 145
QY 72 P-----ITMDNGNIDTELVGTLTGKVT-----GTSTSVNFDADGDPYL 116
DB 146 PLEPSEFRIKQDINDNAPELNGPY---HATVPEMSILGISTVNTATD-ADDPY- 199
QY 117 TFSODGNNHOFITVKYIKSDSRDISPKV-----NGENLVGDVYLAATGSDFFVR 168
DB 200 -----GNSAKLYSIL-EGOPYFSIEPEETAIIKTALPMREAKEEYLVYIQAARD----- 248
QY 169 STGSGKGLAAGKYDATATVVSNOEFMIRAVEQNPQHIEDLEVRQGLGHSSTYLE 228
DB 249 -MGHSGGLSG--TTTLTVLTFD-----VNDPPEFASLVHFSVPE-----DVLG 292
QY 229 ELVOLVKKNIDISIKYDPRKDEVFANRVITDIELLKLILAYFLPBDTLKGGHYDNO 288
DB 229 ELVOLVKKNIDISIKYDPRKDEVFANRVITDIELLKLILAYFLPBDTLKGGHYDNO 288

DB 293 TAIGRVKANDDIGNAQSUSD-----IIDGGTALFEITS-----DAQ 331
QY 289 LONGIRKVEFLESSPTOMELRAFMAVH-----FSLTADRIDDILKIVYDSMNH----- 340
DB 332 AODGVIRLRKPLDEFETKSKYTLKVEANIHIDPREGPGFEDTATVKIVEDADEPPVF 391
QY 341 -----HGDAKRLRELAELTAEK--IYSVIOAEINKH--LSSGTINIHDKSIN 387
DB 392 SSPTYLLEVEHENA--ALNSVIGVYTAADPDITSSPIRESIDRHTDLEROFINADDKIT 449
QY 388 L--MDKNLYGTDEEIEFRAS-AEYKILEKMPOT--TIQVDSSEKKIVSINKDEFLGSENKR 441
DB 450 LATPLDRELVSVMHNISIIATEIRNHQSISRVPALIKVLVDNNAPEFASVEAFLECNCK 509
QY 442 TGAIGLKNKSYSNKDNNELSHF 464
DB 510 PGQV--IQTVSAMDKDDPKNGHF 530

RESULT 13
US-08-332-638-42
Sequence 42, Application US/08332638
Patent No. 5646250
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,638
FILING DATE: 01-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,643
FILING DATE: 17 APR 1992
APPLICATION NUMBER: US/08/049,460
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5646250and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31340
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-332-638-42

Query Match 4.7%; Score 125.5; DB 1; Length 799;
Best Local Similarity 21.3%; Pred. No. 0.078; Indels 117; Gaps 25;
Matches 107; Conservative 80; Mismatches 199; Indels 117; Gaps 25;

QY 19 DDDKHKKISSVATLAFGTATANAADLASTATATL--VEPARTTLT-----YKEGA 71
DB 88 DDDPGSKIKIYILSGGAGIIPQIN--DITGDIHAIKRLDREKAEYTLTAAQVDMETNK 145

QY 72 P-----ITINDGNIDTELLVGTTLGKYT-----GTTSTSVNFTDAGDPMTL 116
DB 146 PLEPSEEFITQYDINDNPEFLNGPY-----HATVPEMSILGTSVYNVATD-ADDPYV- 199
QY 117 TETSQDGNHOFETTKYIGDSRDEISPKV-----NGENLVGDVVLATGSODEFVR 168
DB 200 -----GNSAKLYSTL-EGOPYFSTIEPTALIKTALPMDRPAKEEYLVIAQAD----- 248
QY 169 SIGSKGKLAAGKYTDVATVTVSNQEFMIRAVEONFQHEIDLEKRVQTLGHGSSYLE 228
DB 249 -MGHSGGSLG--TTLVLTILD-----VNDNPKRAQSIXHSVPE-----DYVIG 292
QY 229 ELVQLYKDKNIDISIKYDPRKSEVFANNVITDDIELLKILAVLEPDTILKGHYDQ 288
DB 293 TAIGRVKANDDIGNAQSUSD-----IIDGGTALFEITS-----DAQ 331
QY 289 LONGIRKVEFLESSPTOMELRAFMAVH-----FSLTADRIDDILKIVYDSMNH----- 340
DB 332 AODGVIRLRKPLDEFETKSKYTLKVEANIHIDPREGPGFEDTATVKIVEDADEPPVF 391
QY 341 -----HGDAKRLRELAELTAEK--IYSVIOAEINKH--LSSGTINIHDKSIN 387
DB 392 SSPTYLLEVEHENA--ALNSVIGVYTAADPDITSSPIRESIDRHTDLEROFINADDKIT 449
QY 388 L--MDKNLYGTDEEIEFRAS-AEYKILEKMPOT--TIQVDSSEKKIVSINKDEFLGSENKR 441
DB 450 LATPLDRELVSVMHNISIIATEIRNHQSISRVPALIKVLVDNNAPEFASVEAFLECNCK 509
QY 442 TGAIGLKNKSYSNKDNNELSHF 464
DB 510 PGQV--IQTVSAMDKDDPKNGHF 530

RESULT 14
US-08-755-587-183
Sequence 183, Application US/08755587
Patent No. 6045957
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Selzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh
STATE: NC
COUNTRY: USA

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5405-135
 INFORMATION FOR SEQ ID NO: 183:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1683 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-755-587-183

Query Match 4.7%; Score 124; DB 3; Length 1683;
 Best Local Similarity 18.4%; Pident. No. 0.3; Indels 218; Gaps 34;
 Matches 125; Conservative 120; Mismatches 5

13 SSGHI-DDDKHKMKISSVI-----AIAFGTATNADLSTATATLV 58
 575 ASGLTDENEVEERGFALGTIKANSSALQKVKLF-----SDIENISEIAEV 626
 59 EPARTILT--YKGAFTINDGNID-----TELLVGT--LTIG-----GYKT 97
 627 DPISLSSSKYHDSVASKFIENQXDKSEKNNKCOLLQNNIMETGIFVEENTENYKR 686
 98 GTSTSTVNTDAAGDPYLLFTSQDGNHOTTIVIGKDSKF-----DISKRVNGEN 150
 687 NTEDEMKRYTGASRNSXNLE-XSDGSDSKNDTVYIKHDETDLRPFIDQKNICLKLGQF 745
 151 LVGDDVVLATGODEFYSIGSKGKLAAGKYDAVTYVTSNOEFMIRAVE----- 201
 746 MKGENTQIKBGLSD--LTCLYMAEFTCHGTSKREQLATKTONIKDPTPISTQTA 803
 202 --QNPQHFIEDL-----EKVNEQLTGHSSVLELVQVKNKIDIS----- 242
 804 SGRNIRVSKESLKNVAVNEFQKXTEELNPNFSDLSNSELGINKNMKDISHXETDI 863
 243 IKYDPRDSEY-FANRYIT-----DDIELLKTLAFLPEDTILKGGHNDQNGIKR 295
 864 VKNNILKESPVGTGNQVLTQORECEIEKIK-----EPTLL-GFHTASGRK--VKI 912
 296 VKFELESSPTQWELRAFMVMEFLPADRIDDLKVIYDSNMHGBASRLRE----- 350
 913 AKESLDKVKHL-----FDETEQGVSE-----ITSESHQAKITLKDREACKDG 954
 351 -----ELAEFLAELKIVSYIOAELINKHLSSTGT-----NIHDSINLMDKN----- 392
 955 LELACEVEITAPKCEMQLSLENDKYKSEITVLPQLLSNDLYRQTEMLKTSNSLSLK 1014
 393 --LGYTDEELFKA-----SAEYKILEKMP-----QTTIQ----- 420
 1015 VKHVENVEKETAKSPCTCYTNSYSYVIEENSALAFYTGHSRKTVSSEASLLEAKKWLREG 1074
 421 -VDGSEKKIVSIK-----DFLG-----SENKRTGALGN--LK 449
 1075 IFDDQPRINTAKYCEKEHEDYIGNALYENSNSITENDKNHLSERKODSTYISNSMS 1134
 450 NSYSYNKD-----NNELSHFATTCSDKS--RP-LNDLYSQKTTOLSDITSRPSAIEA 499
 1135 NSYSYHSDFCSDDEVYNDSGLSKNKIDXSIEFLVKNVEDQKNISSEVY-----SAVKE 1190
 500 LNRFTQKIDSY-MORLDDTS 519
 1191 ANTYPTQVNDICVEKLVTNSS 1211

RESULT 15
 US-08-480-604A-24
 Sequence 24, Application US/08480604A
 Patent No. 5736139
 GENERAL INFORMATION:
 APPLICANT: KINK, JOHN A.
 APPLICANT: THALLEY, BRUCE S.
 APPLICANT: PADHYE, NISHA V.
 APPLICANT: FIRCA, JOSEPH R.
 APPLICANT: STAFFORD, DOUGLAS C.
 TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND

TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MEDLEN & CARROLL, LLP
 STREET: 220 MONTGOMERY STREET, SUITE 2200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/480,604A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/422,711
 FILING DATE: 14-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/405,496
 FILING DATE: 16-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/329,154
 FILING DATE: 25-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/429,791
 FILING DATE: 31-OCT-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: INGOLTA, DIANE E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: OPHD-01763
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 23 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-604A-24

Query Match 4.6%; Score 122; DB 1; Length 23;
 Best Local Similarity 79.2%; Pident. No. 0.00081;
 Matches 19; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
 1 MGHNNHHHHHSSGHIIDDDKHM 24
 1 MGHNNHHHHHSSGHI--EGRHK 22

Search completed: August 22, 2000, 16:47:15
 Job time: 1092 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 16:32:34 ; Search time 38.09 Seconds

(without alignments)
846,468 Million cell updates/sec

Title: US-08-699-716a-2

Sequence: 2656 1 MGHNNHHHHHHSSGHIDD.....RFIOKYDSVMORLDDTSGK 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 178050 segs, 61884766 residues

Total number of hits satisfying chosen parameters: 178050

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1646	61.7	326	2 T43594	YOP secretion and
2	1645	61.7	326	2 B33601	lcrv protein - Yersinia
3	1590	59.6	326	2 B37314	regulatory protein
4	850	31.9	170	2 S13008	capsular antigen F
5	164.5	6.2	1252	2 B42771	reticulocyte-bindin
6	150.5	5.6	1116	2 S57382	hypothetical prote
7	149	5.6	2166	2 G70163	hypothetical prote
8	147	5.5	785	2 T50180	cadherin-7 - chick
9	145	5.4	964	2 S57379	MSH2 protein - yea
10	144	5.4	1115	2 T41342	probable coiled-co
11	139.5	5.2	1957	2 T38077	hypothetical colle
12	138.5	5.2	644	2 T46211	hypothetical prote
13	137	5.1	1979	2 C71622	hypothetical prote
14	136	5.1	1312	2 BMBYD1	RAD50 protein - ye
15	136	5.1	2269	2 T28677	trypsin protein -
16	135	5.1	1173	2 T25539	hypothetical prote
17	135	5.1	1530	2 T16927	myosin heavy chain
18	135	5.1	1938	1 MMKWI	myosin heavy chain
19	135	5.1	1938	2 T21193	hypothetical prote
20	134.5	5.0	1467	2 PC1253	Yb protein - yea
21	134	5.0	1939	2 T18372	repeat organellar
22	133.5	5.0	2401	2 T28676	trypsin protein -
23	133	5.0	1039	2 S62509	probable vesicular
24	133	5.0	168	2 T58315	Wtl - human
25	132	5.0	2606	2 T03159	large fragment pro
26	131.5	4.9	547	2 G64851	flagellar hook-ass
27	131	4.9	1803	2 S56894	Yb protein - yea
28	130.5	4.9	1427	2 S22695	reslin - human
29	130	4.9	551	2 T25197	hypothetical prote

30	129	4.8	510	2 G70162	conserved hypothet
31	129	4.8	694	2 G72065	translation elonga
32	129	4.8	694	2 B81602	translation elonga
33	129	4.8	944	2 S26710	spindle pole body
34	129	4.8	1111	1 A42640	kinesin-related pr
35	129	4.8	1465	2 S31262	Yb protein - yea
36	129	4.8	1928	2 S46773	myosin heavy chain
37	128.5	4.8	816	2 S64439	hypothetical prote
38	128.5	4.8	1088	1 P1XRBR	inner layer protei
39	128	4.8	1802	2 S52611	Yb protein - yea
40	127.5	4.8	586	2 C70091	conserved hypothet
41	127.5	4.8	958	2 T20621	hypothetical prote
42	126.5	4.7	1955	3 T30934	myosin-like protei
43	126	4.7	552	2 A71864	ribosomal protein
44	126	4.7	598	2 T11615	hypothetical prote
45	125.5	4.7	348	2 E81292	flagellar P-ring p

ALIGNMENTS

RESULT 1

YOP secretion and targeting control protein - Yersinia pestis plasmid pCD1

C:Species: Yersinia pestis
C:Date: 21-Jan-2000 #sequence-revision 21-Jan-2000 #extl-change 04-Mar-2000

C:Accession: T43594; T42889

R:Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Bruba

J. Bacteriol. 180, 5192-5202, 1998

A:Title: Structural organization of virulence-associated plasmids of Yersinia pestis.

A:Reference number: 222578; M0ID:98422474

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Accession: T43594

A:Molecule type: DNA

A:Residues: 1-326 <HUP>

A:Cross-references: EMBL:AF053946; PIDN:AC62574.1

A:Experimental source: strain KIM

R:Periy, R.D.; Straley, S.C.; Fetherston, J.D.; Rose, D.J.; Gregor, J.; Blattner, F.R

Infect. Immun. 66, 4611-4623, 1998

A:Title: DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersini

A:Reference number: 222273; M0ID:98427122

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Accession: T42889

A:Molecule type: DNA

A:Residues: 1-326 <PER>

A:Cross-references: EMBL:AF074612; PIDN:AC69799.1

A:Experimental source: strain KIM5

C:Genetics:

A:Gene: lcrv

A:Genome: plasmid pCD1

Query Match	61.7%	Score 1646;	DB 2;	Length 326;
Best Local Similarity	99.7%	Pred. No. 1.3e-79;		
Matches 325;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps	0;			
196 MIRAYEONPOHFIDLEKRYEOLTGHSVLEELVOLYDKNDISIKDPKRDSEYFA	255			
1 MIRAYEONPOHFIDLEKRYEOLTGHSVLEELVOLYDKNDISIKDPKRDSEYFA	60			
256 NRVITDIEELKRIILAYFLPEDTILKGYHNOLONGIKRYKEFLSSPNTOWELRAFWA	315			
61 NRVITDIEELKRIILAYFLPEDTILKGYHNOLONGIKRYKEFLSSPNTOWELRAFWA	120			
316 VMHSLRADRIDDIILKYVDSNMHGDASKLAEELAEITAEIKTISVIOAEINKHLSS	375			
121 VMHSLRADRIDDIILKYVDSNMHGDASKLAEELAEITAEIKTISVIOAEINKHLSS	180			
376 SGTINIHDKSINLMDKNLYGTDEIFKASAEYKILEKMPOTTIOVGSEKKIYSINDFL	435			
181 SGTINIHDKSINLMDKNLYGTDEIFKASAEYKILEKMPOTTIOVGSEKKIYSINDFL	240			
436 GSENRKRGALGNLKNYSYKNNKNNELSHFAUTCSDKSRPLNDLVSOQTOLSDITSRFRNS	495			

Db 241 GSENRKTGALGNLKNKSYNKNDELSHFATCSDKSRPLNDLVSOQTQLSDITSRFS 300

QY 496 ATEALNRFIOKIDSVQMRLLDPTSGK 521

Db 301 ATEALNRFIOKIDSVQMRLLDPTSGK 326

RESULT 2

1crv protein - Yersinia pestis

C:Species: Yersinia pestis

C>Date: 17-Jan-1990 #sequence_revision 17-Jan-1990 #text_change 23-Mar-1993

C:Accession: B33601

R:Price, S.B.; Leung, K.Y.; Barre, S.S.; Straley, S.C.

J. Bacteriol. 171, 5646-5653, 1989

A:Title: Molecular analysis of 1crvH, the V antigen operon of Yersinia pestis.

A:Reference number: A33601; MUID:90008806

A:Accession: B33601

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-326 <PRD>

A:Cross-references: GB:M26405

Query Match 61.7%; Score 1645; DB 2; Length 326;

Best Local Similarity 99.4%; Pred. No. 1.5e-79;

Matches 324; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 196 MIRAYQNPQHFIEDELEKRVLEQTLGHGSSVLEELVQVKNKIDISIKYPRKDESEVA 255

Db 1 MIRAYQNPQHFIEDELEKRVLEQTLGHGSSVLEELVQVKNKIDISIKYPRKDESEVA 60

QY 256 NRYITDDIELKKILAYFPEDTILKGHYDNLQNGIKRVKEFLSSPNTQWELRAFA 315

Db 61 NRYITDDIELKKILAYFPEDTILKGHYDNLQNGIKRVKEFLSSPNTQWELRAFA 120

QY 316 VWFSLTADRIDDDILKIVYDSNNHGDARSKLRELAETELTKYVIOAEINKHLS 375

Db 121 VWFSLTADRIDDDILKIVYDSNNHGDARSKLRELAETELTKYVIOAEINKHLS 180

QY 376 SGTINIHDSINIMDKNLGYTDEEIFKASAEKILEKMPOTTIOVDSSEKIVSIKDFL 435

Db 181 SGTINIHDSINIMDKNLGYTDEEIFKASAEKILEKMPOTTIOVDSSEKIVSIKDFL 240

QY 436 GSENRKTGALGNLKNKSYNKNDELSHFATCSDKSRPLNDLVSOQTQLSDITSRFS 495

Db 241 GSENRKTGALGNLKNKSYNKNDELSHFATCSDKSRPLNDLVSOQTQLSDITSRFS 300

QY 496 ATEALNRFIOKIDSVQMRLLDPTSGK 521

Db 301 ATEALNRFIOKIDSVQMRLLDPTSGK 326

RESULT 3

B37314 regulatory protein 1crv - Yersinia pseudotuberculosis

C:Species: Yersinia pseudotuberculosis

C>Date: 30-Jan-1993 #sequence_revision 30-Jan-1993 #text_change 08-Oct-1999

C:Accession: B37314

R:Bergman, T.; Hakansson, S.; Forsberg, A.; Norlander, L.; Macellaro, A.; Baekman, A.;

J. Bacteriol. 173, 1607-1616, 1991

A:Title: Analysis of the V antigen 1crvH-yopBD operon of Yersinia pseudotuberculosis:

A:Reference number: A37314; MUID:91134114

A:Accession: B37314

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-326 <BER>

A:Cross-references: GB:M57893; NID:g155456; PIDN:AAA27645.1; PID:g155458

Query Match 59.6%; Score 1590; DB 2; Length 326;

Best Local Similarity 96.6%; Pred. No. 1.1e-76;

Matches 315; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 196 MIRAYQNPQHFIEDELEKRVLEQTLGHGSSVLEELVQVKNKIDISIKYPRKDESEVA 255

Db 1 MIRAYQNPQHFIEDELEKRVLEQTLGHGSSVLEELVQVKNKIDISIKYPRKDESEVA 60

QY 256 NRYITDDIELKKILAYFPEDTILKGHYDNLQNGIKRVKEFLSSPNTQWELRAFA 315

Db 61 NRYITDDIELKKILAYFPEDTILKGHYDNLQNGIKRVKEFLSSPNTQWELRAFA 120

QY 316 VWFSLTADRIDDDILKIVYDSNNHGDARSKLRELAETELTKYVIOAEINKHLS 375

Db 121 VWFSLTADRIDDDILKIVYDSNNHGDARSKLRELAETELTKYVIOAEINKHLS 180

QY 376 SGTINIHDSINIMDKNLGYTDEEIFKASAEKILEKMPOTTIOVDSSEKIVSIKDFL 435

Db 181 SGTINIHDSINIMDKNLGYTDEEIFKASAEKILEKMPOTTIOVDSSEKIVSIKDFL 240

QY 436 GSENRKTGALGNLKNKSYNKNDELSHFATCSDKSRPLNDLVSOQTQLSDITSRFS 495

Db 241 GSENRKTGALGNLKNKSYNKNDELSHFATCSDKSRPLNDLVSOQTQLSDITSRFS 300

QY 496 ATEALNRFIOKIDSVQMRLLDPTSGK 521

Db 301 ATEALNRFIOKIDSVQMRLLDPTSGK 326

RESULT 4

S13008 capsular antigen F1 precursor - Yersinia pestis plasmid pMT1

C:Species: Yersinia pestis

C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 04-Mar-2000

C:Accession: S13008; T14702; T15015; S23725

R:Galayov, E.E.; Smirnov, O.Y.; Karlishay, A.V.; Volkovoy, K.I.; Denesjuk, A.I.; Nazim

FBS Lett. 277, 230-232, 1990

A:Title: Nucleotide sequence of the Yersinia pestis gene encoding F1 antigen and the

A:Reference number: S13008; MUID:91099503

A:Accession: S13008

A:Molecule type: DNA

A:Residues: 1-170 <GAL>

A:Cross-references: EXBL:X61996; NID:g48620; PIDN:CA443966.1; PID:g48621

R:Hu, P.; Elliott, J.; McCreedy, P.; Showronski, E.; Garnea, J.; Kobayashi, A.; Carra

submitted to the EMBL Data Library, March 1998

A:Description: Structural organization of virulence determinants in three Yersinia pe

A:Reference number: T14702

A:Accession: T14702

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-170 <HUP>

A:Cross-references: EXBL:AF053947; NID:g2996286; PID:g2996338; PIDN:AA013218.1

R:Rindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blatner, F.R.

Infect. Immun. 66, 5731-5742, 1998

A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIMS plas

A:Reference number: Z18268; MUID:99043898

A:Accession: T15015

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-170 <LIN>

A:Cross-references: EXBL:AF074611; NID:g3883003; PID:g3883098; PIDN:AA082758.1

C:Genetics:

A:Gene: caf1

A:Genome: plasmid pMT1

C:Superfamily: Yersinia pestis plasmid pMT1 capsular antigen F1

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-170/Product: capsular antigen F1 #status predicted <MAT>

Query Match 31.9%; Score 850; DB 2; Length 170;

Best Local Similarity 100.0%; Pred. No. 3e-38;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAALFGLTIAANAADLSTATTATATLVEPRITLTITTEGAPITIMONGINTD 83

Db 1 MKRISVIAIALEGTATANAADLTASTATATLVEPARTITLYKECAPITIMDNGNIDT 60
 QY 84 ELLVGLTLLGGYKTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRFDIS 143
 Db 61 ELLVGLTLLGGYKTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRFDIS 120
 QY 144 PKVNGENLVGDDVYVLTATGSDOFFVRSIGSKGKLAAGKTYDATTYVVSNO 193
 Db 121 PKVNGENLVGDDVYVLTATGSDOFFVRSIGSKGKLAAGKTYDATTYVVSNO 170

RESULT 5
 B42771
 reticulocyte-binding protein 2 - Plasmodium vivax (fragment)
 C:Species: Plasmodium vivax
 C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Sep-1997
 C:Accession: B42771
 R:Galinus, M.R.; Medina, C.C.; Ingravallo, P.; Barnwell, J.W.
 Cell 69, 1213-1226, 1992
 A:Title: A reticulocyte-binding protein complex of Plasmodium vivax merozoites.
 A:Reference number: A42771; MUID:92315338
 A:Accession: B42771
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1252 <GAL>
 A:Cross-References: GB:M88098; NID:q160627; PID:q160628
 A:Experimental source: strain Belem, merozoites
 C:Genetics:
 A:Gene: RBP2

Query Match 6.2%; Score 164.5; DB 2; Length 1252;
 Best Local Similarity 20.3%; Pred. No. 0.41;
 Matches 131; Conservative 104; Mismatches 221; Indels 189; Gaps 33;

QY 8 HHHHSSGHIDDD-----KHKRISVIAIALEGTATANAADL----- 47
 Db 402 HHHHSSGHIDDD-----KHKRISVIAIALEGTATANAADL----- 47
 QY 48 -TASTATATL-----VEPARI-----TLVYKCAPITIM 76
 Db 460 ATSENTAEKLEVKVDQSVYVYLNQITTEKRLYTEKRLGIDSTINIGA--LKE 517
 QY 77 DNGNIDTELLVGLTLLG---GYKTTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYI 133
 Db 518 SKGNYEIGLEKLEIGKNNKLVADITKKSINSTVGNFSLFNNF----- 562
 QY 134 GKDSRFDISPKVNG--ENLVGDDVYVLTATGSDOFFVRSIGSKGKLAAGKTYDATTYVVSNO 192
 Db 563 --DLNQYDENKNINDYENKKEIYNEFEGSLNKISENL--RNASENTSDYNSAKTLRLA 618
 QY 193 OEEMIRAY--EONPOHFIELEK-----RVBQLN-----GHGS--S 225
 Db 619 OKKAVVLLKKEEANKYLRVKKVYVESRFTFNKESLIDKINEMIKKQLTVNEGHWKVO 678
 QY 226 VLEELVQVKNKNDISIKYDPRKDEVEFANRYITDDIELKKILAVFLPE--DTLLKGG 283
 Db 679 LVENIKELVDENLSDILKQATGKNEI-----OKITHSTLKNKAKTIL--G 723
 QY 284 HYNOJON-GIKRVKEF-----LESSPPTOMELR--AFMAVHESFLADRID-- 327
 Db 724 HDVTSKAYVGIKTPLEALTELLGDAKLKTAQELKEESKNVNVLEENMKNTNELDVK 783
 QY 328 --DQILKVIYDSMNHG--DARSKLBELELAELKIVIAELINKHLSGGTINIH 393
 Db 784 NIDAKVALEILAHSEIDITKODSKLEMGNOITLYKVL--INQYNNKISSISKE 840
 QY 384 KSLNLDKNLYGYTDEIFKASAEK-----ILEKAPOTTIO-----VDGSEK 426
 Db 841 EAVSVKIGNV-SKKHSELKITCSKSYDNIILAK--QTELDONLANSFTQEKTNNSDS 897
 QY 427 KIYSIK-DLFGSENKRGALGNK--NSYSYNNKNNELSHFATCSGDKRPLNDLVSO 482

Db 898 KLEKITDDESLEKN-----ALKTLEGEVNAALKASDNHE-----HYQSKSEPVNPALSEI 947
 QY 483 TTOLDITSRFNSALEL-----NRFIOKYDSVMORLDDT 518
 Db 948 EKEETDIDS-LMTALDELKAKRTECVSRKYLKIDVTNIEISDDT 991

RESULT 6
 S57382
 hypothetical protein YOL087c - yeast (Saccharomyces cerevisiae)
 N:Alternative names: hypothetical protein 00944
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 12-Dec-1997
 C:Accession: S57382; S66781; S50418
 R:Zumstein, E.; Pearson, B.M.; Kaloogeropoulos, A.; Schweizer, M.
 Yeast 11, 975-986, 1995

A:Title: A 29.425 kb segment on the left arm of yeast chromosome XV contains more than
 A:Reference number: S57374; MUID:96021809
 A:Accession: S57382
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-1116 <ZUM>
 A:Cross-References: EMBL:X83121; NID:q600461; PID:q600471
 R:Zumstein, E.; Pearson, B.M.; Kaloogeropoulos, A.; Schweizer, M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66775
 A:Accession: S66781
 A:Molecule type: DNA
 A:Residues: 1-1116 <ZUM>
 A:Cross-References: EMBL:Z74829; NID:q1419927; PID:q251885; PID:q1419928; MIPS:YOL087
 A:Experimental source: strain S288C
 C:Genetics:
 A:Map position: 15L

Query Match 5.6%; Score 150.5; DB 2; Length 1116;
 Best Local Similarity 18.3%; Pred. No. 1.9;
 Matches 94; Conservative 96; Mismatches 193; Indels 131; Gaps 20;

QY 3 HHHHHHHHSSGHIDDDKHKRISVIAIAL-----FGTATANAADLTASTATA 53
 Db 365 HHHHHHHHEHEDONISTTAKKYGGLIDIALLPNEKLLFECDDNSNVLDLTNNHS 424
 QY 54 TATLVEPARTITLYKEGAPITIMDNGNIDTELLVGLTLLGGYKTTSTSVNFTDAAGD 113
 Db 425 ----VNEGFALTRS-----LTNRHVTENTKGMQMDYSCLELWTFPSSBSGSPD 476
 QY 114 MYLTFTSODGNHQT-----TKVIGKDSRFDISPKVNGENLVGDD 155
 Db 477 IVKRYTSKEILSHMCTVSVKGMLEFKINPKFLKTEVYSALKDQYV--VNNIEINSDE 533
 QY 156 -----VLTATGSDOFFVRSIGSKGKLAAGKTYDATTYVVSNOEEMIRAYEONPOHFIE 209
 Db 534 RYNLGKIVINSLENERI--SYEQKDKLKRKIFSLKKRDLNLSLTDLGYSSESKNK 591
 QY 210 DLEKRVNEOL-----GHGS-----VLEE--LVOLVKNDISIK 244
 Db 592 DKRKRTFKISSLISIGNNSSGTPNSAPATPAVMAETVLEEQPLQASDAID---- 647
 QY 245 YDPRKDSVFANRYITDDIELKKILAVFLPEDTILKGYHNOJONGI--KRVKEFLES 302
 Db 648 -----DSLELVQPLPAKRP-----YFRQSSGSLSRKFRFRST 683
 QY 303 SP-----NTOMELRPAVMAHESLADRIDDLKIVYDSMNHGDAS-----K 347
 Db 684 SGRTATGLTWPEPKILDPDTPHVN--DSAPFOANTNTQOSKADAPESLTMNHPK 739
 QY 348 LREELAELETAE-----LKISVIOAEINKHLSGGTINIDKS-----ILMDKNLYGYD 398
 Db 740 LEOKLSAISQDLPNSNTNKNLRSNSRANSTSTLEGKEKKKPEPMPDLQIOESYKO 799
 QY 399 EEIFKASAEYKILEKMPOTTIOVDSGEKKRIVSIK 432

Db 800 QYNTSSSLKY-LTKRLPYTKI-IRASSCPITRVK 831

RESULT 7

70163

hypothetical protein BR0512 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: G70163

R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Iathigra, R.; White son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vuglt; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943

A:Accession: G70163

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-2166 <L>

A:Cross-references: GB:AE001153; GB:AE00783; NID:92688419; PID:NAAC6876.1; PID:9268842

A:Experimental source: strain B31

Query Match 5.6%; Score 149; DB 2; Length 2166;

Best Local Similarity 21.6%; Pred. No. 5.6;

Matches 111; Conservative 85; Mismatches 154; Indels 164; Gaps 31;

75 IMDNGNIDTELLVGLTGYGKGTSTSVNFTDAGDPMYLT-----TPTS----- 120

890 LSSGKVDLDIDSEVT-----TKIKELKFSIESLESYLEIKIDEFRNGCAIYSEDL 942

121 QDGNHQTIVKIKDSRDP--ISPKV-----NGENLVGD-----DVLATGSG 163

943 QDINMH-----FKETRELEENISKKFAVLNSEEKVEVDSLLQDRIDIASFOANI 996

164 DEFVRSIGSKG--GKLAGKYTDVAVTVSNOEFMIRAYQONQHIEDLEK---VAVE 217

997 DITDLSLVKFNINKEIKNGKINVEITISYRGYSENISKLENEIMHETENLSRRLTDRI 1056

218 QLTGHGSSVLELVQVYKKNIDISIKYPRKDESEVFNRY--ITDIE-----LTKIL 270

1057 SL-----SKGDENELQKKE--SEFVS-KYQVEK-----PELAKYKDLTDEGAENILVKEIE 1106

271 AYFLP-----EDTLKQ--GHVNOQNGIKRVEKFEFLSSPTQWELRAF 314

1107 QYYSRLEEAIDYRRTIDNDIMQAKEREGETNELKNIESSEPL-----NDLYKER--- 1159

315 AVMHFSLAIRIDDDILKVIYDSMHHGDARSKLRELAELTALKTYSVIAQAINKHL 374

1160 -----FKLIESNFEERYSTFLESEG---ATSKIRDE-----YKTL----- 1193

375 SSGTINIHDSKINIMDKRNLGYTDEEIFKASAEKILEKMPQTTIOVDSSEKIV-SIKD 433

1194 TSNDEINQIK-ISEMDN-----PEIIFQKSKDILF--EKELQDKIKD 1224

434 FLGSEKRTGAL-----GNLKNYSYKNKNELSHFATTCSDKSRPL--NDLVSOAKTQOL 486

1235 CYGRINSQFGKIKAGEVENIKN-----HF-DVCIKRVNTLDDIVAYE----- 1277

487 SDIRSNSAIEALNRFQKDKVSMQRLDDTSG 520

1278 -----NEIHKRIDSLKSTESTFSTIEKMLNDKVSQ 1307

RESULT 8

150180

cadherin-7 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000

C:Accession: I50180

R:Nakagawa, S.; Takeichi, M. Development 121, 1321-1332, 1995

A:Title: Neural crest cell-cell adhesion controlled by sequential and subpopulation-s

A:Reference number: I50178; MUID:95309115

A:Accession: I50180

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-785 <NA>

A:Cross-references: GB:D42150; NID:9868000; PID:BA07721.1; PID:9868001

A:Superfamily: cadherin; cadherin repeat homology

F:156-262/Domain: cadherin repeat homology <CDH>

Query Match 5.5%; Score 147; DB 2; Length 785;

Best Local Similarity 23.0%; Pred. No. 1.8;

Matches 115; Conservative 71; Mismatches 201; Indels 114; Gaps 26;

12 HSSGHIDDDKHKMKI--SVYIALFETIATANAADLTAS-----TTATATVEPARIT 64

72 HSD--VKGGSISIKYILSGGASIF--IDENTGDIHAKRKDRDREGAYTTLRQAHR 127

65 LTYEGAP-----TTMDNGNIDTELLVGLTGYK--TGTTSTSVNFTDAGDPMYLT 116

128 LTKPVPESSEFVKIIDIENDEPKFLDGYTAGVPSPVGSVQVATD-ADPTY- 185

117 TFSQDGNHQTIVKIKDSRDPDISPKV-----NGENLVGDVYVLTGSDFFVR 168

186 -----GNSARVYVSI--QGQPYFSEPKTGIIKTALPNDRKADQYLLVIAKDM-- 235

169 SIGSGGKLAGKYTDVAVTVSNOEFMIRAYQONQHIEDLEKVEEQLTGHGSSYLE 228

236 -VGONG--LSG--TTSVYTLID-----VNDNPRFRRSYQVWPE-----SLPLA 278

229 ELVOLVKNIDISIKYPRKDESEVFNRYITDIELKILAYFLPEDTILKGHDNQ 288

279 SVVARIRAADAVG---PNAEMEY--KIVDGGGLGVFKI-----SVDKD 317

289 LQNGIKRVEKFEFLSSPTQWELRAFAMVH-----FSLTARIDDDILKVIYDSMHHGDA 344

318 TQBITITQKELDEAKTSTLRIEAMNHVDPFSLGFSMTYTKIIVEDV----- 371

345 RSKLRELAELTALKTYSVIAQAINKHLSSGTINIHDSIN-----LMDKNLYGYTD 398

372 -----DEPPYFVS--RYSNVVSEAKVGTITIGVAHADDAHSNPRYSIDRN---TD 420

399 -EELFKASAEYKILEKMPQTTIOVDSSEKIVSIKDFLGSSEKRTGALGNLKNYSYKMD 457

421 LERFEINIDANSQVY-----TTAKSLDRETNVAHNITVLAMESQNPQIGGYVAITLID 475

458 NNELSFA-----TTCSDKSRP 474

476 NDNAPEFAMEYETTVCCENAP 496

RESULT 9

557379

MSH2 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein O0935; protein YOL090W

C:Species: Saccharomyces cerevisiae

C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Sep-1999

C:Accession: S57379; S50415; S50782; S56784; S27433

R:Zumstede, E.; Pearson, B.M.; Kalogetopoulos, A.; Schweitzer, M.

Yeast 11, 975-986, 1995

A:Title: A 29,425 kb segment on the left arm of yeast chromosome XV contains more tha

A:Reference number: S57374; MUID:96021609

A:Accession: S57379

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-964 <UNK>

A:Cross-references: EMBL:X83121; NID:9600461; PID:CAAS8189.1; PID:9600468

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994

R:Zumstede, E.; Pearson, B.M.; Kalogetopoulos, A.; Schweitzer, M.

submitted to the EMBL Data Library, December 1994

A:Reference number: S50410

A:Accession: S50415

hypothetical collared-coil protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T38077

C:Conor, R.; Churche, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, April 1996

A:Reference number: 221767

A:Accession: T38077

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1957 <CON>

A:Cross-references: EMBL:270690; PIDN:CAA94624.1; GSPDB:GN00066; SPDB:SPACIF3.06C

A:Experimental source: strain 972h; cosmid c1f3

C:Genetics:

A:gene: SPDB:SPACIF3.06C

A:Map position: 1

Query Match

Best Local Similarity 21.8%; Score 139.5; DB 2; Length 1957;

Matches 103; Conservative 74; Mismatches 161; Indels 135; Gaps 21;

QY 145 KVGNGENLV-----GDDV-----VLATGSDFFVRSIGSGKGLAAGKYDAVT 187

DB 1021 KNGENIASIQTETEKRAENDLQSLVSSSEYENLL--ISSQTNKSLDK----- 1072

QY 188 VYVSNDFMIRAYEONPHIE--DLEKRVESQLTGHSVLEELVOLVDRKNIDISIKY 245

DB 1073 ---TNQ---LKYTEKNQKLEDEKQNRNVELELTSKYGKLGEBNAQ--IKDEL--LALRK 1123

QY 246 DPRDSVFANRYTDDIELLKLIAFLPED-----TIKGGHNDQ-----LNGI 293

DB 1124 KSKQKQDLCAFW--DLKESDALQELTNEKNELVLSQNSNNALVEERSDILNRL 1181

QY 294 KRYKEFLSSPNTQWELRAFMVAFSLTADRIDDILIKY-----IYDSNNHGG 342

DB 1182 SDKKSLSDSDNIVSYRSDLVNDELDTLKKDKDSLSTQYSVCODRDDLDLSLGC 1241

QY 343 DANKLEELAEIT-----AEIKIYVIAQAEINKHLSSSGT 378

DB 1242 ESNKKAVALSRELTQSEIDVPVSEILDNFVFNAGNSELSRLTSLLE--NYLDAFNQ 1299

QY 379 INHDSKINLMQNLVYTDDEIFKASAEKILEKMPQ-----TIQVDSSEKIV--SIKD 433

DB 1300 VNRKHELD---NRLTTTAELTKVAD---LEKIQHEHDDWLIQGDLEKALKDEKN 1352

QY 434 FLGSENRRTGALGNLK-----NSYSYKDNNEISHPATTCSDSRP 474

DB 1353 FLRKEAMTENHSLSECKEETKEKIAELSSRLLEDNOLANKLNQNDH----- 1401

QY 475 LNDLVSKTQQLSDITSRNSAIEALNRFIOKIDSV-----NORLIDTSGK 521

DB 1402 LNDLIRLEKEDVLEKEKELIISLESISLNOROKESSLIDARNEHEHMDTSRK 1454

RESULT 12

T46211

hypothetical protein TBP19.180 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46211

R:Choi, N.; Robert, C.; Bottler, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Sa

submitted to the protein Sequence Database, December 1999

A:Reference number: 223008

A:Accession: T46211

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-644 <CHO>

A:Cross-references: EMBL:AL13315

C:Genetics:

A:Experimental source: cultivar Columbia; BAC clone TBP19

A:Map position: 3

A:introns: 133/1; 373/3; 403/3; 496/3; 566/3

A>Note: TBP19.180

Query Match

Best Local Similarity 19.9%; Score 138.5; DB 2; Length 644;

Matches 113; Conservative 91; Mismatches 211; Indels 153; Gaps 27;

QY 10 HHSSGHIDDDKHKMKISSVIALEGTATANAADLTASTATATLVEPARITLYKE 69

DB 67 HAGVGNSNDKRSAREKASHLALVY-----LODLADSSASELPS--SKRQKN 114

QY 70 GAITTINDG-----NIDELVGLTTLGGYKTGTSTVY-----NTDA 110

DB 115 GNPIDQCDHDEKLVPWKGIIVNIPF-----TAAODGRSAGEGSKLRDEYILRGHPTR 169

QY 111 GDPMY-----LFTSODGN-----NHQFTTVYKGRSDPDISPKV----- 146

DB 170 VRELMTNLGHSATVIEPKNDGMLNGLLFDKAYIVDGHGKDWLKKQPKLGLYGMTA 229

QY 147 -----NGENLVGDVVLVANGSDFFVRSIGSGKGLAAGKYDAVTVYVSNDFMIRAYE 201

DB 230 RADYNGNNIIGENL-----RKTGDLKT--IAELTEEARKQELLY----- 268

QY 202 QNPQHIEDLEK--VAVEQLTGHSSVLEELVOLVDRKNIDISIKY-----DPRKDEY 253

DB 269 QNRLQVLEKMKDKMEIEELCSYKS--EELNOLMEKEKNOCKHYREINAIQERTMSHI 325

QY 254 FANRVITDDIELLKLIAFLPEDITLKGHYNDLONIGIKRYK--EPLD--SSPTQWE 309

DB 326 ---QIVDHEKRLKLLSEKRLLEKCNELAKREYVHNTERKLSIEDLQNAKSSISLE 382

QY 310 LRAFMVAFSLTADRIDDILKVIYDSNNHGDARSKLEELAEITAEIKIYVIAQAEI 369

DB 383 LAA-----MEQKRADEEVKRLAEDQ-----REELHEKIRLERQDRQQAILEYV 427

QY 370 N-----KHLSSGGINHDSKINLMQNLVYTDDEIFKASAEKILEKMPQTTIQ 420

DB 428 EOLKQGLNVRKMAASDQDAEV--VKEVDIIFKDL-----GEKEAQLDLDFKFNQTLIL 478

QY 421 VDSEKKIYVIRKDFSENKRTGALGNLNSYSYNKDNNEISHPATTCSDSRPLNDIYS 480

DB 479 ---RERR---TNDELQAKHE---LVNIMKEWNTNIGVRMGEVLT-----KPEYDAHQ 523

QY 481 QKTQQLSDITSRNSAIEALNRFIOKID 508

DB 524 QKYYC--QVDEDRAYEVILQMEHYLKDSD 550

RESULT 13

C71622

hypothetical protein PFB0145c - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 07-May-1999

C:Accession: C71622

R:Gardner, M.J.; Tetteh, H.; Carnucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.

; Perlea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H

Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600; MUID:99021743

A:Accession: C71622

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1979 <GAR>

A:Cross-references: GB:AE001375; GB:AE001362; NID:g3845105; PID:g3845107; TIGR:PFB014

C:Genetics:

A:Experimental source: clone 3D7

A:gene: PFB0145c

Query Match

Best Local Similarity 20.8%; Score 137; DB 2; Length 1979;

Matches 95; Conservative 71; Mismatches 135; Indels 156; Gaps 20;

```

Qy 120 SODNNHFFTTKVIKGSQSPDISPKYNGNIVGDVPLATSGODPFVRSIGSGGKLA 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 98 NNDNNNN-----NNDNNNNNTSDSEIKTHIKHDELNOLKD-TLNSISLSKTY- 149
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 180 GKTYDAVTVTSNOEFMIRAYEQNPQHFIDLEKVRVEQLTGHGSSVLEELVLYDNKI 239
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 -----NYESKIDELK-----ELKE-VADKNI 170
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 240 DISTIKYDPRDSEVFANRVITDIELLKILAFPLPBDITLKSGHYDNLQNGIKRVEF 299
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 DNNDYENKLEKEFE---VKQKIDMNE-----KENLQREKIDIN---KEKKI 214
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 300 LESSPTOMELRAEAMVHSTLADRIDDDILKYIVDSNNHHGDANSKLEELAEIATL 359
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 NEKEKNIIKKETHNIEKEYLEKKNRETIIEIIDIKH----LEKLIKELKEKKED 270
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 360 KIYSVIOAEINKHLS-----SGTINIHKSINIMDLNKLGYIDEEIFKASBYKLL 411
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 271 E-----NLKKLLSKENVLKELKGCVKERKNETINSLDN-----IIEKKRYKLL 315
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 412 E-KMPQTTIOVDSEK-----KIV 429
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 EYELBEKKQÖIDLNKQDEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKE 375
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 430 SIR-----DLAGSEKNRTGALGNKSNYSYKNNDELSHFPATTSODKSRLNDL 478
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 SIKVREMDEIKERHNFLEMEDQ-----LKDLKNSFV-KNNNOLKVY--KOEIKN-LKTE 425
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 479 VSQKTOLOSDITSRFNSALTEALNFIQKYSVMKRL 515
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 426 LEKKEKELKDIE---NWSKEINKLQNLKEKEKQÖIL 459
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 14

RAD50 protein yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein N0872; protein YN1250w
C:Species: *Saccharomyces cerevisiae*
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 10-Dec-1999
C:Accession: S05808; S63223
R:Alani, E.; Subbiah, S.; Kleckner, N.
Genetics 122, 47-57, 1989
A:Title: The yeast RAD50 gene encodes a predicted 153-kD protein containing a purine nuc
A:Reference number: S05808; MUID:89276917
A:Accession: S05808
A:Molecule type: DNA
A:Residues: 1-1312 <NLA>
A:Cross-references: EMBL:X14814; NID:g4272; PIDN:CAA32919.1; PID:g4273
R:Sen-Gupta, M.; Guelandner, U.; Behnhauer, J.; Fiedler, T.; Hegemann, J.H.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63220
A:Accession: S63223
A:Molecule type: DNA
A:Residues: 1-1312 <SEN>
A:Cross-references: EMBL:L271526; NID:g1302292; PIDN:CAA96157.1; PID:e239641; PID:g1302292
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:RAD50; MIPS:YNL250w
A:Cross-references: SGD:S0005194; MIPS:YNL250w
A:Map position: 14L
C:Superfamily: RAD50 protein
C:Keywords: ATP; coiled coil; DNA repair; meiosis; nucleus
F:177-421/Region: heptad repeats
F:743-995/Region: heptad repeats
F:40/Binding site: ATP (Lys) #status predicted

```

Query Match      5.1e; Score 136; DB.1: Length 1312;
Best Local Similarity 21.58; Pred. No. 14;
Matches 96; Conservative 79; Mismatches 178; Indels 94; Gaps
OY 124 NNHQTTVVIGAGDSADDPDISPKYNNENLVGGDVLATGSGDFPFRSGSGKGLAAKTYT 103
| | | | | : : : | : | | | | |

```

```

Dh 225 NIHQLOKIDQYINNEVEBIEBQLEINEIKEDK--LEFSNDQF--QKILSEVENMLKMTK-- 278
QY 184 DATVTVSNDEFMIRAYEQNPOHEFIEDLEKRYEQOLTGHGSSVYLEELVOLVKDKNIDIST 243
Dh 279 ----LSISDQ--VKRLSNSID--IDLKSPDLOLNLANSKYVLMKNNKNOLEULEDIST 329
QY 244 KYDPRKDESEVAFNRYITDDIELLKILAYFLPBEDTILKGYHDYQLONGIKRYVEKELESS 303
Dh 330 LKDRQSSLOQSNLSNLIIRQOGL-----EAGKRYEYKNNNHLSLSLEAFQHKFOGLSNI 362
QY 304 PNTOW-----ELRPFMAVVMHPSLTADRID-----DILK-VYDYSMN--HHG 342
Dh 383 ENSDMAQYNHEMSOFKAFISODLT-DTIDOFAKDIQLEKTNLSPLIKSITVYDSONEEYK 441
QY 343 DARSKLRELAELTAEIKIYSVIOAEINKHLSGGTINIHDKSINLMDKRLYGTDEIF 402
Dh 442 KDRSKLHJDESELAELKLSF-----KSLSTODSINHELEMLKTYKEKLOQSESENI 493
QY 403 KASAEYKILEMPPQTITTOVQOSEKKIYSIDFQSEMKRT---GALGNLKNVSYNDDN 459
Dh 494 P-----KLNQKIEEKNEMMIILEQIEKFDRIKRTQOQADLYAKGLIKKSINTKLD-- 546
QY 460 ELSHEATCSKSR-----PLN-----DL-----VSOKTQLOLSDITSRFN 494
Dh 547 ELQKTEKLOQDSNRIQOYFLPQLOFEQORADLEMOFKLEINNOKNIAINNNKMHLEDRYT 606
QY 495 SAIEALNRFIQYDSVMOORLLDDSGK 521
Dh 607 NALTYMLN-----TIEKLOLDONOKSK 626

```

RESULT 15

128677
 thopyry protein - Plasmodium yoelli
 C.Species: Plasmodium yoelli
 C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C.Accession: T28677
 R.Keen, J.; Sinha, K.; Brown, K.; Holder, A.
 Mol. Biochem. Parasitol. 65, 171-177, 1994
 A.Title: A gene coding for a high molecular mass thopyry protein of Plasmodium yoelli
 A.Reference number: 220508; MUID:95021522
 A.Accession: T28677
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 17269 <KE>
 A.Cross-references: EMBL:U27838; NID:9457145; PID:9457146; PIDN:AAA21304.1

```

Query Match          5.18; Score 136; DB 2; Length 2269;
Best Local Similarity 22.18; Pred. No. 29;
Matches      89; Conservative 80; Mismatches 141; Indels 92; Gaps 22;

QY  175 GKLAGKTYDATVTVYSNOEFIRAYEONPOHF---IEDLEKVRVEOLGHSVYLEELVQ 232
      |||  |||  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db  611 GKEGIKITTDIKRI--NDDIM-AVSQOIDHINDIOQR---KSESYSEKQJINK 663

QY  233 LVKDNKNIDIS-----IKYPRKDSFVFNRYITD-----DIELLK 268
      |||  |||  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db  664 LEKVSNETELSDNMBVGIKKKQOIIYTKIDKKKNIEELINKLSELSKENTOSLEKYKD 723

QY  269 I-LAY-----FL-----PEDTLKGGHYDNLONGIKRYKKEFLSSPNTOWELR 311
      |||  |||  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db  724 INLSYGONLGNLEFLOIDEKKKAENTIKSMEAYITDDIN--IKKSOEIEEMIDKMDI 782

QY  312 AFMAVMHESLTFADRIDDDILKVIYDSNNHH---GARSKLBEELAEELKIVSIOAE 368
      |||  |||  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db  783 KEMELKLS-----HDDDKCCHDKSKNNHKNISIIYKSSKIIODFSRESDI--NDINK 835

QY  369 INKLHSSSGTINIHDKSINLM---DKNLVGYDEEIFASAEYKILEMPOTIIOVQSE 425
      |||  |||  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db  836 LQKNVSEON---NHSDDINQCLNEVANY-----NILNLNKIKRIDKVRKETSLENK 887

QY  426 KIVSINKDFLGSENKRTGALGNLKNISYNNKNNELSHFATTCSDKS----RPLNDLVS 480

```

Db 888 K--NINDELNSENK---VKKTEGDLSEKCRSKIN---STDDKDIDECIKININVLKR 938
QY 481 QKTTQJSDITSRPSAIE---ALNRF--IQKYDSVMORLLD 516
DB 939 NINNETNITNHFKNABEYKNIIVLSNFNEMADNKSQYILE 980

Search completed: August 22, 2000, 16:50:01
Job time: 1047 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 16:46:59 ; Search time 42.13 Seconds

(without alignments)
383.226 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 266
Sequence: 1 MGHNNHHHHSSGHIDDD.....RFIQKDYDVMQRLDDTSGK 521

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 85661 seqs, 30989116 residues

Total number of hits satisfying chosen parameters: 85661

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1665	61.7	326	1	LCRY_YERPE
2	1590	59.6	326	1	LCRY_YERPE
3	1590	31.9	170	1	LCRY_YERPE
4	164.5	6.2	1251	1	RBP2_PLAVB
5	149	5.6	1433	1	REST_CHICK
6	145	5.4	964	1	MSH2_YEAST
7	139.5	5.2	1357	1	YD6_SCHPO
8	136	5.1	1312	1	RA50_YEAST
9	135	5.1	1338	1	MYSD_CAEEL
10	133.5	5.0	850	1	D7_DICDI
11	133	5.0	1039	1	YAF3_SCHPO
12	131.5	4.9	546	1	FLGK_ECOLI
13	131	4.9	1803	1	YUJ3_YEAST
14	130.5	4.9	1427	1	REST_HUMAN
15	130	4.9	530	1	YGGP_CAEEL
16	129	4.8	694	1	EPG_CHLPN
17	129	4.8	944	1	NUF1_YEAST
18	129	4.8	1111	1	KIP1_YEAST
19	129	4.8	1128	1	MSL_YEAST
20	128.5	4.8	816	1	YG3A_YEAST
21	128.5	4.8	1088	1	RAPO_ROTBR
22	126	4.7	598	1	YAA5_SCHPO
23	125.5	4.7	1088	1	RRO_ROTSL
24	125	4.7	1085	1	CUT7_SCHPO
25	125	4.7	1489	1	YGP0_YEAST
26	124.5	4.7	875	1	ZIP1_YEAST
27	124.5	4.7	1330	1	SMG3_YEAST
28	124	4.7	2663	1	CENE_HUMAN
29	123.5	4.6	799	1	CADB_MOUSE
30	123.5	4.6	1088	1	RRO_ROTBU
31	123	4.6	1639	1	MSPI_PLAFW
32	122	4.6	309	1	YQAP_BACSU
33	122	4.6	800	1	PARC_STRAU

34	122	4.6	1460	1	N159_YEAST	P40477 saccharomyc
35	122	4.6	1858	1	VGNB_CPSNV	P36312 cowpea seve
36	121.5	4.6	539	1	MYX3_HYDAT	P39922 hydra atten
37	121	4.5	1058	1	POL3_DROME	P04323 drosophila
38	121	4.5	1130	1	YU17_CAEEL	O11102 caenorhabdi
39	121	4.5	1312	1	PIP1_DROME	P23455 drosophila
40	120.5	4.5	1125	1	YB62_SCHPO	O14248 schizosacch
41	120.5	4.5	1875	1	MLP1_YEAST	O02455 saccharomyc
42	120.5	4.5	2022	1	ANT1_ONCYO	P21249 onchocerca
43	120.5	4.5	4563	1	APB_HUMAN	P04114 homo saplen
44	120	4.5	556	1	RSL_HELPY	P56008 helicobacte
45	120	4.5	1208	1	YD66_SCHPO	O92351 schizosacch

ALIGNMENTS

```

RESULT 1
LCRY_YERPE STANDARD: PRT; 326 AA.
ID LCRY_YERPE
AC P21206;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN
DE V).
GN LCRY.
OS Yersinia pestis.
OC Plasmid pCD1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM.
RX MEDLINE: 9008806.
RA Price S.B., Leung K.Y., Barve S.S., Straley S.C.;
RT "Molecular analysis of lcrGVH, the V antigen operon of Yersinia
RT pestis."
RL J. Bacteriol. 171:5646-5653(1989).
CC - FUNCTION: POSSIBLY INVOLVED IN CA(2+) REGULATION OF YOP
CC - EXPRESSION, WHICH INCLUDES THE EXPORT PROCESS.
CC - SUBCELLULAR LOCATION: SECRETED.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: M26405; AAA27641.1; ALT_SEQ.
DR PIR: B33601; B33601.
KW Plasmid; Antigen; Virulence.
SQ SEQUENCE 326 AA; 37226 MW; 54FB8209E032F3F4 CRC64;

```

```

Query Match 61.7%; Score 1645; DB 1; Length 326;
Best local similarity 99.4%; Pred. No. 4.1e-78;
Matches 324; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 196 MIRAYQONOHFFDEDEKRVDELGTGSSVLEETVOLYKDRNIDISIKYDRKQSEVRA 255
DB 1 MIRAYQONOHFFDEDEKRVDELGTGSSVLEETVOLYKDRNIDISIKYDRKQSEVRA 60
OY 256 NRYITDIELKKILAFYLPEDTILKGYHNDOLONGIKRVEFELESSPNTOMELAFNA 315
DB 61 NRYITDIELKKILAFYLPEDTILKGYHNDOLONGIKRVEFELESSPNTOMELAFNA 120
OY 316 VAFSLTADRIDDDILKAVYDSNNHNGARSKURELAELTLELKITYSYQAEINKHLS 375
DB 121 VAFSLTADRIDDDILKAVYDSNNHNGARSKURELAELTLELKITYSYQAEINKHLS 180

```

```

QY 376 SGTINHDKSNIMADKMLYGTDEEIFKASAEYKILEKMPOTTIOVDGSEKKIYSIDFL 435
DB 181 SGTINHDKSNIMADKMLYGTDEEIFKASAEYKILEKMPOTTIOVDGSEKKIYSIDFL 240
QY 436 GSEKRTGALGNLKNKSNYKNDNLSHFATTCSDKSRPLNDLVSOKTTOLSDITSRFS 495
DB 241 GSEKRTGALGNLKNKSNYKNDNLSHFATTCSDKSRPLNDLVSOKTTOLSDITSRFS 300
QY 496 AIEALNRFIOKYDSVMORLLDPTSGK 521
DB 301 AIEALNRFIOKYDSVMORLLDPTSGK 326

RESULT 2
LCRV_YERPS STANDARD; PRT; 326 AA.
ID LCRV_YERPS STANDARD; PRT; 326 AA.
AC P23994;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN
DE V).
GN LCRV.
OS Yersinia pseudotuberculosis.
OC Bacterioplasmid pIB1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:YPIII;
RX MEDLINE: 91154114.
RA Bergman T., Hakansson S., Forsberg A., Norlander L., Macellaro A.,
RA Baeckman A., Boellin I., Woll-Waltz H.;
RT "Analysis of the v antigen lcrGVH-yopBD operon of Yersinia
RT pseudotuberculosis: evidence for a regulatory role of LcrH and
RT LcrV."
RL J. Bacteriol. 173:1607-1616(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97162308.
RA Roggenkamp A., Gelger A.M., Lettritz L., Kessler A., Heesemann J.;
RT "Passive immunity to infection with Yersinia spp. mediated by anti-
RT recombinant V antigen is dependent on polymorphism of v antigen."
RL Infect. Immun. 65:446-451(1997).
CC -1- FUNCTION: INVOLVED IN CA(2+) REGULATION OF YOP EXPRESSION, WHICH
CC INCLUDES THE EXPORT PROCESS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M57893; AAA27645.1; -
CC EMBL: X96802; CAA65594.1; -
CC DR PIR: B37314; B37314.
CC DR Plasmid: Antigen: Virulence.
CC KM Plasmid: Antigen: Virulence.
CC SO SEQUENCE 326 AA; 37336 MW; 2FD9455DAFF48C06 CRC64;

Query Match 59.6%; Score 1590; DB 1; Length 326;
Best Local Similarity 96.6%; Pred. No. 2,7e-75;
Matches 315; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 196 MIRAYEONPOHFEDELEKAYVEOLTGHSYVLELVOLVMDKNIDISIKIDPKRDSVFVA 255
DB 1 MIRAYEONPOHFEDELEKAYVEOLTGHSYVLELVOLVMDKNIDISIKIDPKRDSVFVA 60
QY 256 NRYITDIELKTLIAFLPEDRTILKGHYDNOLONGIKRYKELESPTQWELRAFMA 315

```

```

DB 61 NRYITDIELKTLIAFLPEDRTILKGHYDNOLONGIKRYKELESPTQWELRAFMA 120
QY 316 VMHPSLTADRIDDILKYIVDSNMHHGDARSKLREELAEYKILEKMPOTTIOVDGSEKKIYSIDFL 375
DB 121 VIHESLTADRIDDILKYIVDSNMHHGDARSKLREELAEYKILEKMPOTTIOVDGSEKKIYSIDFL 180
QY 376 SGTINHDKSNIMADKMLYGTDEEIFKASAEYKILEKMPOTTIOVDGSEKKIYSIDFL 435
DB 181 SGTINHDKSNIMADKMLYGTDEEIFKASAEYKILEKMPOTTIOVDGSEKKIYSIDFL 240
QY 436 GSEKRTGALGNLKNKSNYKNDNLSHFATTCSDKSRPLNDLVSOKTTOLSDITSRFS 495
DB 241 GSEKRTGALGNLKNKSNYKNDNLSHFATTCSDKSRPLNDLVSOKTTOLSDITSRFS 300
QY 496 AIEALNRFIOKYDSVMORLLDPTSGK 521
DB 301 AIEALNRFIOKYDSVMORLLDPTSGK 326

RESULT 3
CAFL_YERPE STANDARD; PRT; 170 AA.
ID CAFL_YERPE STANDARD; PRT; 170 AA.
AC P26948;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE FL CAPSULE ANTIGEN PRECURSOR.
GN CAFL.
OS Yersinia pestis.
OC Bacterioplasmid pPfr.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91099503.
RA Galinov E.E., Smirnov O.Y., Karlish A.V., Volkovoy K.I.,
RA Dalvadnyak S.M., Zaytsov I.V., Rudtsov K.S., Adiraman V.M.,
RA Denesjuk A.I., Nazimov I.V., Rudtsov K.S., Adiraman V.M.;
RT "Nucleotide sequence of the Yersinia pestis gene encoding Fl antigen
RT and the primary structure of the protein. Putative T and B cell
RT epitopes."
RL FEBS Lett. 277:230-237(1990).
CC -1- SUBCELLULAR LOCATION: CAPSULE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X61996; CAA43966.1; -
CC DR PIR: S13008; S13008.
CC DR Plasmid: Signal.
CC KM Plasmid: Signal.
CC FT SIGNAL 1 21
CC FT CHAIN 22 170
CC FT DOMAIN 100 150
CC FT POTENTIAL.
CC SO SEQUENCE 170 AA; 17666 MW; 9AC87796A0BA67D9 CRC64;

Query Match 31.9%; Score 850; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 1,1e-37;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 MKRISSYIALALGTLATANADLSTATATLVEPARITLYKKGAPITMDNGNIDT 83
DB 1 MKRISSYIALALGTLATANADLSTATATLVEPARITLYKKGAPITMDNGNIDT 60
QY 84 ELVGLTLLGSGYKGTGTSVNFDAAGDPMYLFTLSOGGNNQFTTKYIGKSRPFDIS 143
DB 61 ELVGLTLLGSGYKGTGTSVNFDAAGDPMYLFTLSOGGNNQFTTKYIGKSRPFDIS 120

```

```

OY 144 PRVGENLVGDVVLATGSGDFVFRSISGSGKGLAAGRYTDAVTYSNO 193
DB 121 PRVGENLVGDVVLATGSGDFVFRSISGSGKGLAAGRYTDAVTYSNO 170

RESULT 4
RBP2_PLAVB STANDARD; PRT; 1251 AA.
AC 000799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RETICULOCYTE BINDING PROTEIN 2 (FRAGMENT).
GN RBP2.
OS Plasmodium vivax (strain Belen).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92315338.
RA Galsinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
RT merozoites."
RL Cell 69:1213-1226(1992).
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (PROBABLE).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M88098; AAA29744.1;
KW Malaria; Receptor; Membrane.
FT NON TER 1 1
FT NON TER 1 1
SO SEQUENCE 1251 AA; 143741 MW; 54BA51C740A4C572 CRC64;

Query Match 6.2%; Score 164.5; DB 1; Length 1251;
Best Local Similarity 20.3%; Pred. No. 0.24;
Matches 131; Conservative 104; Mismatches 221; Indels 189; Gaps 33;

OY 8 HHHHSSGHIDDD-----KHKKISVYAIATLFTIATANAAD----- 47
DB 402 HNYDNAEDELNDKSTNAKVLLETMLSEYKH--NLSEITINIKGGEKITYSKAKDIMOKIX 459
OY 48 -TASTATATL-----VEPARL-----ILTYEGAPITIM 76
DB 460 AISENTAKTEKRYVDOSNYVNLQITERNLIVTEKNRLNGIDSTITNIEGA--LKE 517
OY 77 DNGNDITDELIVGTLG---GYKGTSTSVFTDAAGDPMWLEFSTODGNHQTITVI 133
DB 518 SKGNVEIFLEKLEIGKRNKLIKADITKKSINSTVGNFSLNNF----- 562
OY 134 GKSDHDPDISPVNG-ENLVGDVVLATGSGDFVFRSISGSGKGLAAGRYTDAVTYSN 192
DB 563 --DLNQYFNKNINDEYKMGETINEPFGSLNKISENL--RNASENTSYNNAKTLRLA 618
OY 193 QEFMIRAY--EQNPQHEFLEDEKV-----RYEQLT---GHGS--S 225
DB 619 QKEKVNLLNKEEANKYLYADYKVSFFIFENMKESLDKINEMIKKEQLTVNEGHNAYQ 678
OY 226 VLEEVLQYKNDISIKYPRKDSVPANVITDDIELLKILAYLPE--DTILKNG 283
DB 679 IVENKEIVDENLSDILKQATGKNEITL-----OKITHSTLTKNAKTL--G 723
OY 284 HYNDLQON-GIKRVEF-----LESSPNTQWEIR--AFMAVMHFSLTADRID-- 327

```

```

DB 724 HVDTSATYVGKIKTPELALTELLGDALKTRAOELKESKNNVYLETENMSKNTNELDVHK 783
OY 328 --DDILKVIYDSNNHHC--DARSKLREELAEFLAEIKIYVTOAINKHSSGCTINIH 383
DB 784 NIDQAYVALEILLAHSEIDITKQKSSKLIEMGNQIYLVK--INQYKKNKISSISKE 840
OY 384 KSINLMDKNTLYGTDEITFNASAEYK-----ILEKPPQTTIQ-----VDGSEK 426
DB 841 EAVSVKIGNV-SKHSSELSKITSCKSDSYDNILAEK--QTELQWLRNSFTQEKNTNNSDS 897
OY 427 KIVSIRK-DELGSEKRRKGALGNLK--NSYSYNNDNNELSFATCSCKSRPLNDIVSOK 482
DB 898 KLEKIKKIDFESLKN-----ALKTEGEVNAKASDNEH-----HVOSKEPVPAUSEI 947
OY 483 TTQSLDITSRNSAIEAL-----NRFQKDSVMQRLDDT 518
DB 948 EKEETDIDS-LNTALDELKKGRCFVSRKLIINDYTKESIDT 991

RESULT 5
REST_CHICK STANDARD; PRT; 1433 AA.
AC 042184; 042228; 057563; 057564;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE RESTIN (CYTOPLASMIC LINKER PROTEIN-170) (CLIP-170).
GN RSV.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98137792.
RA Griparic L., Volosky J.M., Keller T.C. III;
RT "Cloning and expression of chicken CLIP-170 and restin isoforms."
RT Gene 206:195-208(1998).
RN [2]
RP SEQUENCE OF 17-1139 FROM N.A. (CLIP-170(11) AND CLIP-170(11+35)).
RC TISSUE=PECTORALIS MUSCLE;
RA Griparic L., Keller T.C. III;
RT "Identification and expression of two novel CLIP-170/Restin isoforms
RT expressed predominantly in muscle."
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC CYTOSKELETON (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST FOUR ISOFORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 1 CAP-GLY DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF014012; AAC60344.1;
DR EMBL; AF020764; AAC60345.1;
DR EMBL; AF045650; AAC03547.1;
DR EMBL; AF045651; AAC03548.1;
DR PFMW; PF01302; CAP_GLY. 2.
DR PROSITE; PS00845; CAP_GLY. 1; 2.
KW Cytoskeleton; Microtubules; coiled coil; Alternative splicing.
FT DOMAIN 79 121 CAP-GLY.
FT DOMAIN 144 207 SER-RICH.
FT DOMAIN 235 277 CAP-GLY.
FT DOMAIN 305 332 SER-RICH.
FT DOMAIN 351 1353 COILED COIL (POTENTIAL).

```

FT DOMAIN 1414 1427 CCHC-BOX.
 FT VARSPLIC 458 492 MISSING (IN SHORT ISOBROM).
 FT VARSPLIC 458 492 TOTKLEHAIKEBOSLEKTRAKDLQLEDETR ->
 FT VARSPLIC 803 803 RRRQISEDEBN (IN ISOFORM CLIP-170(11)).
 FT VARSPLIC 458 458 S -> GGSSKVS (IN ISOFORM CLIP-170(11)).
 FT VARSPLIC 458 458 T -> RKROISEDEBN (IN ISOFORM CLIP-170(11+35)).
 FT CONFLICT 309 309 K -> R (IN AAC03547).
 FT CONFLICT 440 440 E -> V (IN AAC03548).
 FT CONFLICT 440 440
 SQ SEQUENCE 1433 AA; 161026 MW; 5631CE8683498E23 CRC64;

Query Match 5.6%; Score 149; DB 1; Length 1433;
 Best Local Similarity 20.3%; Pred. No. 1.8; 212; Indels 140; Gaps 28;
 Matches 115; Conservative 100; Mismatches 212; Indels 140; Gaps 28;

QY 38 TATANAADULTATATATLVEPARI-----TLTYKE-----GAP 72
 DB 140 TAAASATPTSTSTASASASPALLSPGIPKTSPLAKENSTPSQFNSLKTSAGSV 199
 QY 73 ITIMDNGNI---DELLVGLTLGKTKGTTSTSVNF---TD-AAGDPMTLTFTSODGNN 125
 DB 200 SMISEAGSLKKGKREIKIGRVLVG---GTRKAGVRFELGETDFKGEWCEVEDEPLGKN 256
 QY 126 HOFFTVIGKSDHDPDISP-----KVGENDLYGDVVL---158
 DB 257 ---DGAIVAG--TRVFOCOPRYGLFAVHVHYTKIGFPSTTPAKKTTVRKVATPALKNS 311
 QY 159 ---ANGSDFEVRISGSGKGL---AAGKYTDAVTVVSNQEFMIRAYEQNPQHFIE 209
 DB 312 PSASLSLSVSVSVSKSPRGTGLTETSSRYARKISGTALQ---ALKRQOHIEG 367
 QY 210 ---DLEVRVQULGHSSVLELVOLVKNN---IDISIKYPRKDEVFANRVIT 260
 DB 368 LAEROLEAEVAKATSHVGEIQELA-LVRDHDHRYVEMEMKQOLRAVWZADR---423
 QY 261 DDELLKRTI-----LAVFLPEDTLKGYHNDOLONGIK--RYKERELESPTOME 309
 DB 424 EKVELLNOLBEERKVEDLOFRKEESTIKG---DETOTKLEHAIKE-LEOS-----473
 QY 310 LRAFAVMEFSLTADRIDDILKVIYDSNMHNGDASKLREELAEITAEIKIYSVIOAEI 369
 DB 474 -----LFEKTRAKDLQLEDETRVATVSE-----KSRIMELEHDLALAKREVAEL 519
 QY 370 NKHLSSSGINIHDKSINIMDKVLGYTTDEITKASAEKILEKAPOTTIOWDGSKKIV 429
 DB 520 KRLSSKRIHDVDTLSLQLE---ISSLOEKMAAKREHO--REMSLKEKESSEALR 575
 QY 430 SIKDFSENKRTGALG-NLKNSTY--NKDNNELSHFATTCSKSRPLNDLVQKTTOLS 487
 DB 576 KEKTLTASENRGENESLKTLDANKENSDVIELM-----KSK-LESALASHOQAME 629
 QY 488 DITSRENSAIEALNRFIOKYDSVMQRL 514
 DB 630 ELKVSFNKGVCAQTAEFAELKTOMEKV 656

RESULT 6
 MSH2_YEAST STANDARD; PRT; 964 AA.

AC P25847; Q12423;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA MISMATCH REPAIR PROTEIN MSH2.
 GN MSH2 OR YOL090W OR O0935.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 93093456.
 RA Reenan R.A.G., Kolodner R.D.;

RT * Isolation and characterization of two Saccharomyces cerevisiae genes
 RT encoding homologs of the bacterial Hexa and Muts mismatch repair
 RT proteins.";
 RL Genetics 132:963-973(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / FY1679;
 RX MEDLINE; 96021609;
 RA Zumstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
 RT "A 29.425 kb segment on the left arm of yeast chromosome XV contains
 RT more than twice as many unknown as known open reading frames.";
 RL Yeast 11:975-986(1995).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE; 93093457;
 RA Reenan R.A.G., Kolodner R.D.;
 RT "Characterization of insertion mutations in the Saccharomyces
 RT cerevisiae MSH1 and MSH2 genes: evidence for separate mitochondrial
 RT and nuclear functions.";
 RL Genetics 132:975-985(1992).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE; 96413311.
 RA Alan E.;
 RT "The Saccharomyces cerevisiae Msh2 and Msh6 proteins form a complex
 RT that specifically binds to duplex oligonucleotides containing
 RT mismatched DNA base pairs.";
 RL Mol. Cell. Biol. 16:5604-5615(1996).
 CC -1- FUNCTION: INVOLVED IN POST-REPLICATIVE DNA-MISMATCH REPAIR. BINDS
 CC TO MISMATCH-CONTAINING DNA.
 CC -1- SUBUNIT: HETERODIMER OF MSH2 AND MSH6.
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M84170; AAA34802.1;
 DR EMBL; X83121; CAAS8189.1;
 DR EMBL; 274832; CAA99102.1;
 DR PIR; S27433; S27433.
 DR SGD; L0001190; MSH2.
 DR PFAM; PF00488; MUTS.C; 1.
 DR PFAM; PF01624; MUTS.N; 1.
 DR PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; 1.
 KW DNA repair; ATP-binding; DNA-binding.
 FT NP_BIND 688 695 ATP (POTENTIAL).
 FT CONFLICT 957 964 KYRALL -> EIKRSCCYN (IN REF. 1).
 SQ SEQUENCE 964 AA; 108884 MW; 43FED8A640138A84 CRC64;

Query Match 5.4%; Score 145; DB 1; Length 964;
 Best Local Similarity 19.5%; Pred. No. 1.7;
 Matches 116; Conservative 98; Mismatches 212; Indels 168; Gaps 27;

QY 47 LTAATATATLVEPARITLTYEGAPITIMNG-----NIDEL 85
 DB 84 VIVSLQVATLTLKLCULDLDGYK---VEIYDGMWLISASPGNIEOVNEIMNMNIDISI 139
 QY 86 LV-----GLTLGKGYRTGTTSTSVNFTDAAGDPMTLTFTS-----120
 DB 140 IIASLKYVMNSODSGCIIIVAITTAAYVGLDLYDNVEVSNLESLILOGKCKLYOD 199
 QY 121 ---ODGNHOTTTRVIG-----KSDHFDISRVNGENLVGDVYLVATGSQDF 166
 DB 200 LTNNSNSNAEMQKVINVIDRCGVVTLTKNSFESEKDYELDTKLGDIALSLT-PQKYS 258
 QY 167 VASIGSKGKLA-----AGKYTDAVTVVSNQEFM-----IRAYEQNPQHF 207

```

Db 259 KLSMGACNALLGYQLSEDOVGKYE--LVEHKLEFEMKIDASAIKALNLEFPGQPNP 315
QY 208 IEDLEKVEAEQTLGHSS-VLEELVOLYKDKNIDISIKY-----DPRKSEVFANR--- 257
Db 316 FGS--NNLAVSGFTSAGNCGKXTSLFQLNHCKNTAGVLLNWKLPPLNIDINKRDL 374
QY 258 -VITDIELLKILAYLP-----EDTLKGHYDNOIQRKRYKEFLESSPMTQ 307
Db 375 VDYLIDQIELKQMLTSEYLPMPIDIRLTKLNKRGNEEDVL-----KYGOSKRIPET- 428
QY 308 WELRAFMAVHFSLTADIDDLKVIYDSNMHGDASKRLRELAELTAEIKITS----- 363
Db 429 --VOVFTEFLDDSPFEPEVNELVRSWMLAPLSHAEPLSKF-EZEVETTVDAIAYENNE 485
QY 364 -VIOAEINK-----HLSSGTINIH-DKSNLMDKNLYGY----- 396
Db 486 FMKVEFEHELCKIKSKDLTLDLHSHSDSAEDLGFDPKRLLENHHLHGRCMLRTR 545
QY 397 TDEEIFKASAEYKLEKMPQTTIOVGSEKKIVSIKDELGSE-NRRTGALG---NLKNS 451
Db 546 NDAKELRKHKKYIELSTVAKAGIFSTKQLKSANETNIIQKRYDQOSALVREIINTLT 605
QY 452 YS--YKNNNELSH-----FATTSCKSRP-----LNDLYSQKTTQSDITSR 492
Db 606 YTPVEKLSLVLAHLDAVJASFAHTSSYAPIPYIRPKLHPMDSERTHTL-ISSR 657

RESULT 7
YD86_SCHPO STANDARD; PRT; 1957 AA.
ID YD86_SCHPO Q10411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYDROTHERMAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN SPAC1F3.06C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
NC Schizosaccharomycetaceae; Schizosaccharomycetes.
RC [1]
RP SEQUENCE FROM N.A.
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z70690; CAA94624.1; -
RW Hypothetical protein.
SO SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

Query Match 5.2%; Score 139.5; DB 1; Length 1957;
Best Local Similarity 21.8%; Pred. No. 8.4;
Matches 103; Conservative 74; Mismatches 161; Indels 135; Gaps 21;

QY 145 KVNENLV-----GNDV-----VLATGSODFFVRISIGSGKGLAGKTTDAVT 187
Db 1021 KNNGNINISLQTEIKRAENDLDQSKLSVSEYENLL--ISSQTNKSLSDK----- 1072
QY 188 VYVSGEEMIRAYEQNPQHFIE--DLEKRYEQLTGHSSVLELYVOLYKDKNIDISIKY 245
Db 1073 ---TNG---LKYEKNVQKLDEKQKRVNVELELTSKYKGLGEENAO--TKDEL--LALRK 1123
QY 246 DPRKSEVFANRYTDDIELLKILAYLPED-----TLKGGHYDNO-----LQNGI 293
Db 1124 KSKKQHDICANFV--DDLKESDALEQLTNRKNEILVLSLEQSNNEALVEERSDLANRL 1181

```

```

QY 294 KRYVELESSPNTQWELRAFMAVHFSLTADRIDDLIKV-----IYDSNMHNG 342
Db 1182 SDMKRSLSDSNVYSVIRSDLYRVNDELDTLAKRKDLSLTPYSVQCDRODLDSLNGCE 1241
QY 343 DARSKLEELAEI-----AEIKIYVIOAEINKHLSSGT 378
Db 1242 ESFNKYAVSLNELCTKSEIDVPVSEIIDDNFVFNAGNSELSRLTVLSLE--NYLDAFNQ 1299
QY 379 INIDKSNIMDKNLGYTDEEIFKASAEYKILEKMPQTT--TIYDGSKKIV-SIKD 433
Db 1300 VNFKMEID---NRLLTTTDEFTKVAD---LEKLQHEHDDWLIQGGDLERALKDSEKN 1352
QY 434 FLGSENNKRTGALGNLK-----NSYSYKDNNELSHFATTSCKSRP 474
Db 1353 FLRKEAENTENHSLBEKEETKKEIAELSRLEDNOLATKRLKNQD----- 1401
QY 475 LNDLYSQKTTQSDITSRFSNLAIEALNLFIOKYDSV-----MORLLDDTSRK 521
Db 1402 INOEIRLEKEDVLEKESLILISESLSNQOKESSLLDAKNELEHMLDDTSRK 1454

RESULT 8
RASO_YEAST STANDARD; PRT; 1312 AA.
ID RASO_YEAST P12753;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNA REPAIR PROTEIN RASD50 (153 KDA PROTEIN).
GN RASD50 OR YNL250W OR N0872.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
NC Saccharomycetaceae; Saccharomycetes.
RC [1]
RP SEQUENCE FROM N.A.
RA STRAIN-RB821;
RX MEDLINE; 89276917.
RA Alani E., Subbiah S., Kleckner N.;
RT "The yeast RASD50 gene encodes a predicted 153-kD protein containing a
RT purine nucleotide-binding domain and two large heptad-repeat
RT regions."
RL Genetics 122:47-57(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE; 97377992.
RA Sen-Gupta M., Gueldeener U., Beinbauer J.D., Fiedler T.A.;
RL "Sequence analysis of the 33 kb long region between ORC5 and SUI1
RL from the left arm of chromosome XIV from Saccharomyces cerevisiae."
RL Yeast 13:849-860(1997).
CC -1- FUNCTION: IS REQUIRED DURING MEIOSIS IN THE PROPHASE FOR
CC CHROMOSOME SYNAPSIS AND HOMOLOGOUS RECOMBINATION. DURING
CC VEGETATIVE GROWTH IT IS NECESSARY FOR DNA REPAIR. THE PROTEIN MAY
CC BE INVOLVED IN SEARCH FOR HOMOLOGY BETWEEN INTERACTING DNA
CC MOLECULES OR CHROMOSOMES. COULD BE A 5'-3' EXONUCLEASE.
CC WORK IN COMPLEX WITH MRE11.
CC -1- SIMILARITY: TO OTHER HEPTAD REPEAT CONTAINING PROTEINS, SUCH AS
CC MYOSINS AND TROPOMYOSINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14814; CAA32919.1; -
DR EMBL; X96722; CAA65494.1; -
DR EMBL; Z71526; CAA96157.1; -
DR PIR; S05808; BMYIDL.

```

DR SGD: I0001570; RAD50.
 KW Repeat: Heptad repeat pattern; DNA repair; ATP-binding; Coiled coil;
 FT Melosits.
 FT NP_BIND 34 41 ATP (POTENTIAL).
 FT DOMAIN 177 421 COILED COIL (35 HEPTADS).
 FT DOMAIN 743 995 COILED COIL (37 HEPTADS).
 SQ SEQUENCE 1312 AA; 152568 MM; 58A0AA173AC5677E CRC64;

Query Match
 Best Local Similarity 5.1%; Score 136; DB 1; Length 1312;
 Matches 96; Conservative 79; Mismatches 178; Indels 94; Gaps 19;

QY 124 NNQFTTKYIGKSRDPDIPKNGENVDDVYATGSDDFVRSIGSGKLAGKRT 183
 DB NHHQLOPKTDQYEEVSELESQNETTEKSDK--LFRSNDP--OKILSKVENIKNTK-- 278
 QY 184 DAVTVSNQEFMIRAEQNPQHFIEDLEKRYEOLGHSVLEELVOLVKKNIDIS1 243
 DB ---LISDQ---VKRLNSID--IIDLSKPDQNLNANFSKYLMDKNQRLDLETDISS 329
 QY 244 KYDPKRDSEFANRVITDDIELLKILAYFLPEDTILKGHYDNOLONGIKRYKEFLSS 303
 DB 330 LKROQSLSLSNSLIRQOEL-----EAGKETEYKNNHLSLKEAFQHKFGLSNI 382
 QY 304 PNTQW-----ELRAFYAVVHFSITADRID-----DILK-VYVDSMN--HIG 342
 DB 383 ENSDAQVHNSQFKAFTISQDIT-PTIDQFANDIQLEKTNLSDLIKSIVDSQNLNEYK 441
 QY 343 DASKRLRELAELTALTKIYVIOAEINKHSSGGINIHDKSINIMDKNLVYGTDEIF 402
 DB 442 KDRSKIHSEELAEKLSF-----KSLSTQDSINHELEMLKYYKEKLSQWESENII 493
 QY 403 KASAEYKILKMPQTTIQVDGSEKKIVSINDFLGSENKRT--GALGNLKNSSYNKDN 459
 DB 494 P-----KLNQKIEKKNNEMITLENQIEKFQDRIMKTNQADLYAKILKKSINTKLD-- 546
 QY 460 ELSHFATTSQDKSR-----PLN-----DL-----VQKQTQLSDITSRN 494
 DB 547 ELQKITEKQNDRIKQVFPFLQEFPRADLEMDFOKLFIMQKRNIAINNKMHEDLRRT 606
 QY 495 SAIEALNRFIOKYDVMORLLDDTSGK 521
 DB 607 NALYNLN-----TIKDLQDQMSK 626

RESULT 9
 ID MYSD.CAEEL STANDARD; PRT; 1938 AA.
 AC P02567; 019674;
 DT 21-JUL-1986 (rel. 01, Created)
 DT 01-OCT-1989 (rel. 12, Last sequence update)
 DT 15-FEB-2000 (rel. 39, Last annotation update)
 DE MYOSIN HEAVY CHAIN D (MHC D).
 GN MYO-1 OR R06C7.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE: 89178677.
 RA Ddb N.J., Matuyama I.N., Krause M., Karn J.;
 RT "Sequence analysis of the complete Caenorhabditis elegans myosin
 RT heavy chain gene family.";
 RT J. Mol. Biol. 205:603-613(1989).
 RP [2]
 RP SEQUENCE OF 34-1795 FROM N.A.
 RX MEDLINE: 83273600.
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the Caenorhabditis elegans unc-54
 RT myosin heavy chain gene are not separated by introns.";
 RT Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).

RN [3]
 RP SEQUENCE OF 115-365 AND 1492-1763 FROM N.A.
 RX MEDLINE: 85201409.
 RA Karn J., Ddb N.J., Miller D.M.;
 RT "Cloning nematode myosin genes.";
 RT Cell Muscle Motil. 6:185-237(1985).
 RP [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Gardner A., McMurray A.;
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MUSCLE CONTRACTION
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS OF THE MYOFIBRILS.
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.
 CC ELEGANS.
 CC -1- SIMILARITY: THE GLOBULAR HEAD AA SEQUENCE SHOWS A HIGH DEGREE OF
 CC SIMILARITY WITH THE GLOBULAR HEAD SEQUENCE OF MUSCLE & NONMUSCLE
 CC HEAVY CHAINS. BY CONTRAST THE ROD SEQUENCE IS LESS CONSERVED, BUT
 CC THE PERIODICITIES OF HYDROPHOBIC & CHARGED RESIDUES, WHICH DICTATE
 CC THE ALPHA-HELICAL COILED-COIL STRUCTURE, ARE CONSERVED.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X08065; CA930854.1; -
 CC EMBL: M37232; AAA28119.1; -
 CC EMBL: M37234; AAA28120.1; -
 CC EMBL: Z71265; CA95848.1; -
 CC EMBL: Z71261; CA95848.1; JOINED.
 CC EMBL: Z71261; CA95806.1; -
 CC EMBL: Z71266; CA95806.1; JOINED.
 CC PIR: S02772; MWKML.
 CC HSSP: P08799; 1MND.
 CC WORMPEP: R06C7.10; CE06253.
 CC PFAM: PF01576; Myosin_tail; 1.
 CC PFAM: PF00063; Myosin_head; 1.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC MYOSIN: muscle protein; Coiled coil; Thick filament; Actin-binding;
 CC ATP-binding; Methylation; Alkylation; Heptad repeat pattern;
 CC Multigene family.
 CC
 CC GLOBULAR HEAD (S1).
 CC ROD-LIKE TAIL (S2 AND LM DOMAINS).
 CC ALPHA-HELICAL TAILPIECE (SHORT S2).
 CC LIGHT MEROMYOSIN (LM).
 CC COILED COIL (POTENTIAL).
 CC ATP (BY SIMILARITY).
 CC ACTIN-BINDING.
 CC METHYLATION (TRI-).
 CC ALKYLATION (SH-1).
 CC ALKYLATION (SH-2).
 CC F -> E (IN REF. 2).
 CC A -> R (IN REF. 2).
 CC V -> D (IN REF. 4).
 CC DV -> GD (IN REF. 2).

FT CONFLICT 391 391 V -> D (IN REF. 4).
 FT CONFLICT 408 408 W -> N (IN REF. 2).
 FT CONFLICT 474 474 Q -> G (IN REF. 2).
 FT CONFLICT 577 577 L -> F (IN REF. 4).
 FT CONFLICT 681 681 I -> N (IN REF. 4).
 FT CONFLICT 1373 1373 S -> D (IN REF. 2).
 FT CONFLICT 1659 1659 E -> Q (IN REF. 3).
 SQ SEQUENCE 1938 AA; 223255 MW; 387399C8F63A4CF4 CRC64;

Query Match 5.1%; Score 135; DB 1; Length 1938;
 Best Local Similarity 19.8%; Pred. No. 14;
 Matches 100; Conservative 89; Mismatches 213; Indels 102; Gaps 19;

QY 76 MNGNID-----TELVGTLTGGYKGTSTSVNFTDAADPMY--LFTSODGN-NHQ 127
 DB 1287 LSNDSNDLAROYELEHILATINRAKTAFSSQVLEAKKAEDLHROEHAKCKLHEH 1346
 QY 128 -----FTTKVIGKDSRDPDIS-----PKVNGENVGD-----D 155
 DB 1347 LPOCHELLEETQINGKDIOGORLSRINSEISQWKAARYEGELVGELEELKRNQNRVD 1406
 QY 156 VYLTATSGDFEYFRSISGSKGKLAGYTDAYVTYSNOEFMTAYQONQH--IEDLKR 213
 DB 1407 LDEALSAONKAYISLEKANGKLA--ETEDARSVDYRHLTVLASLEKORAFDKIYDDMK 1464
 QY 214 VREOGLTHGSSVLEELVOLVKDNIDISIKYDPKDSVEFANRV---ITDDIELKKI 269
 DB 1465 RAYVDI-----QKEIDATTRDSRNTSTEVFKIRSSMDNLSEQIEFLRRE 1508
 QY 270 LAYFLPE-----DTILKGHYDNOQNGIKRVEKEFLESSPNTQWELRAFV---AVMHF 319
 DB 1509 NKIFSGEIDINEQITQGRTYQEVHKSVRLEQEKDELQHALDEAALAEESKVLRL 1568
 QY 320 SILADIDDDILKVIYDSNNHGDARSKLREELAELEAKITSYQAEINKHLSGGIT 379
 DB 1569 QLEVOQIRSEIERKIOEKEEFENRKNHORALESIOASLETEAKKAELARAKKLEET- 1627
 QY 380 NHDKSNINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVGSKTIYSID-FLGSE 438
 DB 1628 DINOELALDHANKANAYDQKNGK-----KLPOYKELQGVDDDEORRRELENTLAEE 1682
 QY 439 NRTGALGNLKNSSYKNKNLSEHPATCSDKSR-----PLNDLYSOKTITDITSRF 493
 DB 1683 KRLALAL-----SSEEDLAH-RIEASDKHKKQLEIEQAEKLSMTELLGN----- 1726
 QY 494 NSAIELNRFIOKYDSVMQRLDD 517
 DB 1727 NAALSAMKRRKVENVOIARNELE 1750

RESULT 10

DB D7_DICDI STANDARD; PRT; 850 AA.
 AC P54682;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CAMP-INDUCIBLE PRESPORE PROTEIN D7 PRECURSOR.
 GN D7.
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Dictyostelida; Dictyostelium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX3;
 RX MEDLINE; 95080502.
 RA Agarwal A., Sloger M.S., Oyama M., Blumberg D.D.;
 RT "Analysis of a novel cyclic Amp inducible prespore gene in
 RT Dictyostelium discoideum: evidence for different patterns of CAMP
 RT regulation.";
 RL Differentiation 57:151-162(1994).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY IN THE PRESPORE CELLS;
 CC -1- INDUCTION: BY CAMP.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; U25143; AAA73514.1; .
 DK DICTYDB; DD02038; .
 KM Sporulation; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 850 CAMP-INDUCIBLE PRESPORE PROTEIN D7.
 FT DOMAIN 470 475 POLY-GLN.
 FT DOMAIN 555 568 POLY-ASN.
 FT DOMAIN 728 738 POLY-GLN.
 SQ SEQUENCE 850 AA; 95343 MW; 13BA634CCE7AA502 CRC64;

Query Match 5.0%; Score 133.5; DB 1; Length 850;
 Best Local Similarity 19.4%; Pred. No. 5.7;
 Matches 107; Conservative 95; Mismatches 202; Indels 147; Gaps 27;

QY 13 SSGHIDDDDKHKKISSVIALALFETATANAD-----LTASTATATLVEPARITL 65
 DB 156 TGTGDSGDSKTYNE-----AIIFFSKVSTTDQESIIQGVATLNDLSGATITGLGVSS 210
 QY 66 TYREGAPITIMDNGINDELLVGLTGLGYKTGTSTSVNFTDAADPMYLTFTSODGN 125
 DB 211 TAKVGQIT--NGRAQGVYITGDNNTGVGCAVTASAVANTGE--FLGSGRTGSS 265
 QY 126 HQFTTKVIGKDSRDPDISKRVNGENVGDVY-----LATGSDDFVRSISGSKLA--- 178
 DB 266 SAGT--VG-----NVISDSTYSIGKIASNGNSJSETITGIGTLAHTF 307
 QY 179 AKRYTDAYVT-----VSNQEFMIRAEQNPQHFIEDLEKRVREOLTG 221
 DB 308 AG--TDSVGVTFPHITKTFNLIAGKRFSDQYIDKSGTIPS---QDNEIKRRLGSA 362
 QY 222 HSSSVLEF---LVOLVKDNIDISIKYDPKDSVEFANRVITDDIELKTLAYFLPED 277
 DB 363 H--QOLQESPAIYQSMKSEDL-----KNIDDEVIRNTLAKEMQIORENDQIGQONED 413
 QY 278 -----TILKGHYDNOQNGIKRVEKEFLESSPNTQ--WELAFAYAMHFSILADRI 326
 DB 414 KQQLIDLQNRPEGLYKNQ-----QDLKQ--ERRANQDELINELMLQEDQEYELLDDL 466
 QY 327 DDDILKVIYDSNNHGDARSKLREELAELEAKITSYQAEINKHLSGGITINIDKSI 386
 DB 467 YDE-----QOOQOPKAVSKQQLQF--QOINSPEIDQIYOLN-HLNDPPDDVTHNDOT 514
 QY 387 NLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVGSSEKKIYSIKDFLGSNKRRTGALG 446
 DB 515 EELKDDVYNFNDQIINNQGFEENV--EEFPD-----LNDANDNEQYN 555
 QY 447 NKNSSYKNKNLSEHPATCSDKSRPLNDLYSOKTITDITSRFNSAIALNRFIOK 506
 DB 556 NNNNNNNNNNNN---FVY---DKSK-----KSAQOVEI-----ALENERLYLOE 594
 QY 507 YDSVMQRLDD 517
 DB 595 VEDAPERLYEE 605

RESULT 11

DB YAF3_SCHPO STANDARD; PRT; 1039 AA.
 AC 009857;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HYPOTHETICAL 118.6 KDA PROTEIN C29E6.03C IN CHROMOSOME 1.


```

Db 344 VAIGATVVDASAVLATDYKRISFDNNQWOTRLASNTTFVTP-DANGKVAFGLE----- 397
Oy 218 OLTGSGSSVEELVOLVMDKNIDISIKYDPRKDEVFANRVITDIELLKLILAEFLPED 277
Db 398 -LTFGTGTRAVD-----SFTLKPSDAIVNMDVLTIDEK-----INASEED 439
Oy 278 TILKGGHVNOLONGIKRYKEFLSSPNTQWELRAFMVHPSLTADRIDDDILKAVYDS 337
Db 440 ---AGSDSN--RNGQALLD--LQSNKRYVGAKSP----- 467
Oy 338 NMHEDASKLEELAEELTAELKITYVIOAEI-----NKHLSGCTINIHDKSINLMDKN 392
Db 468 ---NDAVASIVSDIGNKTATLTKTSSATQGNVTVQLSNOQOISIG-VNIDEYGNLORPO 522
Oy 393 LYGTDEIEFKAS 405
Db 523 OYTLNAOVLQTA 535

RESULT 13
YL3_YEAST STANDARD: PRT: 1803 AA.
ID YL3_YEAST
AC P47024; P87192;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DE 15-FEB-2000 (Rel. 39, Last annotation update)
DE TRANSPONON TY4 207.7 KDA HYPOTHETICAL PROTEIN.
GN TY4B OR YL1113W OR J0780.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
  Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE: 97103775;
RA Czaplepruch C., Korde E., Pujol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including URA2 (5' end), TRK1, PDS2,
RT SP10, GCN14, REP1, PHO86, NCA3, ASF1, CCT7, GEF3, two tRNA genes,
RT three remnant delta elements and a Ty4 transposon.";
RL Yeast 12:1471-1474(1996).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: Z49389; CA889409.1;
RM Transposable element; Hypothetical protein.
RM SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;
SQ

```

Query Match 4.9% Score 131; DB 1; Length 1803;
 Best Local Similarity 19.3%; Pred. No. 21;
 Matches 123; Conservative 103; Mismatches 232; Indels 178; Gaps 31;

```

Oy 13 SSGHIDDD--DKRMKISSVIAIALFGTATANADL-----TASTATATL 57
Db 993 NNSHDEKDYDEKVRHPIPIOEKLVGSKNTIKINDENKISIRKSKNIGSLINCLASC 1052
Oy 58 VEPARTLTYKGCAPITTMNDGNIDTELLVGLTLAGYTF-----GTSTSVNFTDA--- 109
Db 1053 VITDESITNKDES-----MHNAPKELIOEOLKTKNHTSEPPKESIGTNYKFRNTNNE 1106
Oy 110 ---AGDPYVLTFTSODGNH---OFTTVIGKDSRDFISPKVNGENLVGDDVVLATGS 162
Db 1107 ISLKTGSDTSLPKITLESINNHSNDSTYNKVEFEKENHHPEDIEDIVMSQDTMESNC 1166
Oy 163 ODFVFRISGSKGKLAGKYTD-----AVTVVS---NOE 194

```

```

Db 1167 QD-----GNNIKELKVTDKNVPTDNGTNSPRLEQNTIAGSGSPQVTVNKSAPLKE 1217
Oy 195 FMI-----RAYEONPOHFIEDLE-----KYRVEOLTGSSVLEELVOLK----- 235
Db 1218 FSSLNKKRRKRHRDKNNSLTSTELERDKKRRKRVKILPDMETVASPKIRAIYNEAI 1277
Oy 236 DKNDISIKYDPR-----KDSEVFANRVITDIELLKLILAEFLPEDTILKGGHY 285
Db 1278 SKNPDLKKEHYKQAYHKELOLNDKMYF-----DVDV--KTSRSRIPNLVPTPTI 1328
Oy 286 DNQLONGIKRYKEFL---SSPNTQWELRAFMVHPSLTADRIDDDILKY---IVDSMN 339
Db 1329 FTKKRNQYKKAIVCRGDTQSPDT-----YSVTTESINHNHKKIFLIAMNRN 1377
Oy 340 ---HGGDASKLEEL-----AELTAELKITYVIOAEI-----EINKHLS-SGTI 379
Db 1378 MEMKTLIDINHAFLYAKLEELIYIPPHDRRCVVKLNKALYGLKQSPKEMNDHLQYLNGI 1437
Oy 380 NINHDS-----INLMDKNLY--GYTDEIEFKASAEYKI---LEKMPQT-TIQVDSEKKI 428
Db 1438 GLKDSYTPGLYQTEDKMLTAVYDDCVIAASNDQRDEIRINKLSNFEKIKITGLIDD 1497
Oy 429 VSKIDFLGE---NKRIGALGNLKNKY-----SYNKD-----NNELSHFATTCSDKSRP 474
Db 1498 VLDFTILGMDLYVNRRLGTIDLTLSKFINRMDKRYNEELKIRKSSIPHMSTY---KIDP 1554
Oy 475 LNDVYSQKTTQSLDTSRFSNAIEFLN--RFIOKYD 508
Db 1555 KKDVLQMSSEEFROGVLKLOQLGELNRYRHRCRYD 1590

RESULT 14
REST_HUMAN STANDARD: PRT: 1427 AA.
ID REST_HUMAN
AC P30622;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE RESTIN (CYTOLASTIC LINKER PROTEIN-170 ALPHA-2) (CLIP-170) (REED-
DE STERNBERG INTERMEDIATE FILAMENT ASSOCIATED PROTEIN).
GN RSN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PERIPHERAL BLOOD MONOCYTES;
RX MEDLINE: 92289675;
RA Blibe G., Delabie J., Brueggen J., Richener H., Asselbergs F.A.M.,
RA Carletti N., Sorg C., Odink K., Tarsay L., Wiesendanger W.,
RA de Wolf-Peters C., Shipman R.;
RT "Restin: a novel intermediate filament-associated protein highly
RT expressed in the Reed-Sternberg cells of Hodgkin's disease.";
RL EMBO J. 11:2103-2113(1992).
RN [2]
RP SEQUENCE FROM N.A.
RM MEDLINE: 92405160.
RX Piere P., Schuel J., Rickard J.E., Krels T.E.;
RA "CLIP-170 links endocytic vesicles to microtubules.";
RL Cell 70:887-900(1992).
CC -----
CC CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC CYTOSKELETON.
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERNBERG CELLS
CC OF HODGKIN'S DISEASE.
CC -1- SIMILARITY: CONTAINS 1 GAP-GLY DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X64838; CAA6050.1;
DR EMBL: M97501; AAA55693.1;
DR PIR: S22695; S22695.

DR MIM: 179838;
DR PFM: PFO1302; CAP_GLY; 2.
DR PROSITE: PS00845; CAP_GLY; 1; 2.
DR CYTOSKELETON: Microtubules; Coiled coil; Alternative splicing.

FT DOMAIN 78 120 CAP_GLY.
FT DOMAIN 143 204 SER-RICH.
FT DOMAIN 232 274 CAP_GLY.
FT DOMAIN 304 331 SER-RICH.
FT DOMAIN 350 1342 COILED COIL (POTENTIAL).
FT DOMAIN 1408 1421 CCHC-BOX.
FT VARPPLIC 457 481 MISSING (IN SHORT ISOTOPIC).
FT CONFLICT 1069 1069 D->E (IN REF. 2).
SQ SEQUENCE 1427 AA; 160989 MW; 0A4F16DD94254E8 CRC64;

Query Match 4.9%; Score 130.5; DB 1; Length 1427;
Best Local Similarity 20.3%; Pred. No. 16;
Matches 75; Conservative 74; Mismatches 134; Indels 87; Gaps 17;

OY 167 VRSIGSGKGL--AGKYTDVAVTVVNOEFMIRAYEONPOHFE-----DLEKVEEQL 219
DB 326 VSSRPRTGILFTSSRYARKISGTALQE---ALKEKQHQHLEQLAEDELRAYEAKA 381
OY 220 TGGSSVLEVLQVADKN---IDISIKYDPKDESEVFNRTITDIELKTI----- 269
DB 382 TSHVGEIEELA-LARDGHOVLELAKMDQRTWEADR--EKVELNLDEEKKR 437
OY 270 ---LAFLEPEDITLKGHYDNOLONGIK--RYKEFLSSPNTOMELRAFVAVHFSILAD 324
DB 438 VEDLOPVEESTTKG---DETQTKLEHARKIE-LEQS-----LLEKTKAD 481
OY 325 RIDDDILKIVDSNNHNGDARSKRLRELAELIYVIOAEINKHLSSTGTTINIDK 384
DB 482 KLORELEDRVAVTSE-----KSRIMLEKEDLALRQVAVELRRLESNNKPADVDV 533
OY 385 SINDMKNTLGYTDEIFRASAETKLEKMPOTTIOVDSSEKIVSIKDFLSEKRTGA 444
DB 534 SLSL-----OBI-----SSLOEKLEVTTRD--HOREITLKEHFGARE----- 571
OY 445 LGLNKSYSYKNDNELSHPATGSDKSRPLNDVSOQTTOJSDISRFSALEALNRFI 504
DB 572 -----THQEKRLTYATEKLSKENSLSKLEHANKENDVIALWMSKLETA---I 620
OY 505 OKYDVMQRL 514
DB 621 ASHQAMEEL 630

RESULT 15
Y6P_CAEEL
ID YG6P_CAEEL STANDARD; PRT; 530 AA.
AC P90970; P91831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-FEB-2000 (Rel. 35, Last annotation update)
DE HYPOTHEICAL 60.7 KDA PROTEIN T23G11.8 IN CHROMOSOME 1.
GN T23G11.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Gardner A.;

RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: 281130; CAB03421.1;
DR EMBL: 284712; CAB03421.1; JOINED.
DR EMBL: 284712; CAB06546.1;
DR EMBL: 281130; CAB06546.1; JOINED.
DR WORMPEP: T23G11.8; CE14094.
DR Hypothetical protein; Transmembrane.
FT TRANSMEM 505 525 POTENTIAL.
SQ SEQUENCE 530 AA; 60716 MW; AB4376D10C891FA3 CRC64;

Query Match 4.9%; Score 130; DB 1; Length 530;
Best Local Similarity 21.4%; Pred. No. 4.7;
Matches 110; Conservative 84; Mismatches 211; Indels 108; Gaps 24;

OY 40 ATANAD-----LRASTATATLVEPARITLTYSKCAPITINDGNIDELLVTLT 91
DB 7 AIONMADSAPRTKRSSTASVSATIIIMPEHVTTEFGRRAP--SLPSRAQSD---YTGAGM 62
OY 92 LGG--YKGTGTSVVFDTAAGDPVTLFTSODGNNHQTTVICKDSDFISRVKNE 149
DB 63 RGRNRNDOSSSTHQPGVTAASEMTVINSKEVAN-----AKPDEDTISQSSVGR 114
OY 150 NLVGDVVLATGSGQDFVRSIGSGKGLAAGKYTDVAVTVVNOEFMIRAYEONPOHFE- 208
DB 115 SVAS-----IRITSE--FAGNDSSPNKSQLFAKDSQISVLKTKLSEVEKFRSQDYIE 168
OY 209 -----EDLEK-YRVEDLTGHGSSVLEL-----VOLYKDKNIDISIKYDPKDESEVFN 256
DB 169 MRAEKEMLEKRVENKVSSEHSDSLQELKARAKADQKEKAVE--ECNMHRKRIYGL 225
OY 257 RYTDDIELKILAYFLEPEDITLKGHYDNOLONGIKRVKEFLSSPNTOMELRAFNAV 316
DB 226 EELRANVEQLR--LAKFNLENKKEPEDEYKNAQK-ILTAKELVESLSKEGIGSSDRP 282
OY 317 MHESLTAD---RIDDDILKIVDSNNHNGDARSK--LRELAELIYVIOAEINKHLSSTGTTINIDK 368
DB 283 VHL-LQAEVEELRVERDLKADLES-----AQLVYTLRSMEELEAQRLOLSQ 335
OY 369 INKHLSSSTINIDKSTLMKNTL-----YGYDEELFRKSAEY--KLEKMPOTTIQ 420
DB 336 KTHLEKQT--WDSTIGLNEKVECSRIENEFKOEKKRGGDOHOSMLEK----- 385
OY 421 VDGESEKIVSINDFLGSEKRTGALGNLKNYSYKNDNELSHPATGSDKSRPLNDV 480
DB 386 -----ENELKRTISDLRSKMRDDQTOE-----DDGSSQLTDL 420
OY 481 OKTTOJSDITSR---FNSAIEALNRFIOKYDSV 510
DB 421 OKOQLEDEVLRNNQVNLVRLERLQKATNETSI 453

Search completed: August 22, 2000, 16:51:53
Job time: 294 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 16:46:39 ; Search time 46.06 seconds
(without alignments)
784.261 Million cell updates/sec

Title: US-08-699-716A-2

Sequence: 1 MGHNNHHHHSSGHIDDD.....RFIOKDYVMQRLDPTSGK 521

Scoring table: BLOSUM62
Gapop: 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SPTREMBL_12:*
2: sp-archaea:*
3: sp-bacteria:*
4: sp-fungi:*
5: sp-human:*
6: sp-invertebrate:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1646	61.7	326 2	068697 yersinia pe
2	1557	58.4	324 2	087495 yersinia en
3	1552.5	58.2	334 2	056853 yersinia en
4	1549.5	58.1	334 2	056896 yersinia en
5	1537	57.7	325 2	056880 yersinia en
6	1532	57.5	324 2	056882 yersinia en
7	1531	57.4	324 2	056846 yersinia en
8	1527	57.3	324 2	056879 yersinia en
9	1480	18.4	294 2	030527 pseudomonas
10	150.5	5.6	1116 3	099247 saccharomyc
11	149	5.6	2166 2	051465 borrelia bu
12	147	5.5	785 13	090763 gallus gall
13	144	5.4	1115 3	094488 schizosacch
14	139.5	5.2	1928 5	097291 plasmodium
15	139	5.2	924 5	015738 dictyostell
16	137	5.1	1979 5	096133 plasmodium
17	136	5.1	603 4	09Y607 homo sapien
18	136	5.1	2269 5	026223 plasmodium
19	135	5.1	674 2	09XD51 moraxella c

20	135	5.1	1173 5	P92199 caenorhabdi
21	135	5.1	1510 5	Q22699 caenorhabdi
22	134	5.0	1939 5	Q25662 plasmodium
23	133.5	5.0	601 5	Q76379 echinococcu
24	133.5	5.0	2401 5	Q26216 plasmodium
25	133	5.0	795 5	Q9Y1J5 dictyostell
26	132	5.0	16256	Q16256 homo sapien
27	132	4.9	2606 12	Q36414 alcelaphine
28	131	4.9	1803 3	P87192 saccharomyc
29	129	4.8	510 2	Q51457 borrelia bu
30	129	4.8	694 2	Q92802 chlamydia p
31	129	4.8	819 5	Q9XYP7 dirosophila
32	128.5	4.8	1730 10	Q04454 arabidopsis
33	128	4.8	1542 5	Q97299 plasmodium
34	127.5	4.8	586 2	Q45598 bacillus su
35	127.5	4.8	958 5	Q9XTM0 caenorhabdi
36	127	4.8	576 2	Q54407 moraxella c
37	126.5	4.7	1955 5	Q61308 parascaris
38	126	4.7	552 2	Q9XKE6 helicobacte
39	125.5	4.7	799 11	Q54800 rattus norv
40	125	4.7	1173 3	Q13313 schizosacch
41	124.5	4.7	761 13	P79793 gallus gall
42	124.5	4.7	1231 5	Q24098 dirosophila
43	124	4.7	523 2	Q31304 borrelia bu
44	124	4.7	858 1	Q47968 haloferax s
45	124	4.7	1418 3	Q12267 saccharomyc

ALIGNMENTS

RESULT 1
ID 068697; PRELIMINARY; PRT: 326 AA.
DC 068697; 01-AUG-1998 (TREMREL. 07, Created)
DT 01-AUG-1998 (TREMREL. 07, last sequence update)
DR 01-MAY-1999 (TREMREL. 10, last annotation update)
DE V ANTIGEN.
GN LCRV.
OS Yersinia pestis.
OG Plasmid pCD1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM;
RA HU P., ELLIOTT J., MCCREADY P., SKOMRONSKI E., GARNES J.,
RA KOBAYASHI A., CARRANO A.V., BRUBAKER R., GARCIA E.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBD databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM5;
RX MEDLINE: 98427122.
RA PERRY R.D., STRALEY S.C., FETHERSTON J.D., ROSE D.J., GREGOR J.,
RT "DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of
RT Yersinia pestis KIM5.";
RL Infect. Immun. 66:4611-4623(1998).
DR EMBL: AF053946; AAC62574.1; -;
DR EMBL: AF074612; AAC63799.1; -;
KW Plasmid.
SQ SEQUENCE 326 AA; 37240 MW; 443BDEDC CRC32;

Query Match 61.7%; Score 1646; DB 2; Length 326;
Best local similarity 99.7%; Pred. No. 8.9e+80;
Matches 325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 196 MIRAYQNPQHFLIEDLEKVRVQLTGHGSSVLELVQLVKRNKIDISIKYPRKQSEVFA 255
DB 1 MIRAYQNPQHFLIEDLEKVRVQLTGHGSSVLELVQLVKRNKIDISIKYPRKQSEVFA 60

QY 256 NRVTTDDIELLKKILAFPEPDTILKGGHYDNOQNGIKRKEFLESSPNTQWELRAFMA 315
 DB 61 NRVTTDDIELLKKILAFPEPDTILKGGHYDNOQNGIKRKEFLESSPNTQWELRAFMA 120
 QY 316 VMEHSLADRIDDDILKVIYDSNMHHGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 375
 DB 121 VMEHSLADRIDDDILKVIYDSNMHHGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 180
 QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPQTTIOVDSSEKKIYSINDFL 435
 DB 181 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPQTTIOVDSSEKKIYSINDFL 240
 QY 436 GSEKRRGALGNLKNKSYNKNDELSHFATTCSDKSRPLNDVSOXTTQSDITSRPN 495
 DB 241 GSEKRRGALGNLKNKSYNKNDELSHFATTCSDKSRPLNDVSOXTTQSDITSRPN 300
 QY 496 AIEALNRFIOKYSVMORLLDDTSGK 521
 DB 301 AIEALNRFIOKYSVMORLLDDTSGK 326

RESULT 2
 087495 PRELIMINARY; PRT: 324 AA.
 ID 087495
 AC 087495
 DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE LCRV PROTEIN.
 OS Versinia enterocolitica.
 OC Plasmid pTV, and plasmid pYve227.
 OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Versinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-W22703.
 RA IRIARIE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
 RT "Complete sequence of the Versinia enterocolitica pTV plasmid.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-W22703.
 RA IRIARIE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
 RT "Detailed genetic map of the pYve227 plasmid of Versinia
 enterocolitica serotype O:9.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF102990; AAD16815.1; -
 KW Plasmid
 SQ SEQUENCE 324 AA: 37296 MW: 7C9A652F CRC32;

Query Match 58.4%; Score 1557; DB 2; Length 324;
 Best Local Similarity 95.4%; Pred. No. 4.3e-75;
 Matches 308; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 196 MIRAEONPOHFIEDLEKRVVEQLTGHGSSVLEELVOLVYDKNIDISIKYDPKRDSEVFA 255
 DB 1 MIRAEONPOHFIEDLEKRVVEQLTGHGSSVLEELVOLVYDKNIDISIKYDPKRDSEVFA 60
 QY 256 NRVTTDDIELLKKILAFPEPDTILKGGHYDNOQNGIKRKEFLESSPNTQWELRAFMA 315
 DB 61 NRVTTDDIELLKKILAFPEPDTILKGGHYDNOQNGIKRKEFLESSPNTQWELRAFMA 120
 QY 316 VMEHSLADRIDDDILKVIYDSNMHHGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 375
 DB 121 VMEHSLADRIDDDILKVIYDSNMHHGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 180
 QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPQTTIOVDSSEKKIYSINDFL 435
 DB 181 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPQTTIOVDSSEKKIYSINDFL 240
 QY 436 GSEKRRGALGNLKNKSYNKNDELSHFATTCSDKSRPLNDVSOXTTQSDITSRPN 495

DB 241 ESEKRRGALGNLKNKSYNKNDELSHFATTCSDKSRPLNDVSOXTTQSDITSRPN 300
 QY 496 AIEALNRFIOKYSVMORLLDDT 518
 DB 301 AIEALNRFIOKYSVMORLLDDT 323

RESULT 3
 056853 PRELIMINARY; PRT: 334 AA.
 ID 056853
 AC 056853
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE V ANTIGEN (FRAGMENT).
 OS Versinia enterocolitica.
 OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Versinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WA-314;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESEMAN J.;
 RT "Passive immunity to infection with Versinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96796; CAA65588.1; -
 FT NON TER 334
 SQ SEQUENCE 334 AA: 38241 MW: 71229360 CRC32;

Query Match 58.2%; Score 1552.5; DB 2; Length 334;
 Best Local Similarity 93.1%; Pred. No. 7.7e-75;
 Matches 309; Conservative 8; Mismatches 6; Indels 9; Gaps 1;

QY 196 MIRAEONPOHFIEDLEKRVVEQLTGHGSSVLEELVOLVYDKNIDISIKYDPKRDSEVFA 255
 DB 1 MIRAEONPOHFIEDLEKRVVEQLTGHGSSVLEELVOLVYDKNIDISIKYDPKRDSEVFA 60
 QY 256 NRVTTDDIELLKKILAFPEPDTILKGGHYDNOQNGIKRKEFLESSPNTQWELRAFMA 315
 DB 61 NRVTTDDIELLKKILAFPEPDTILKGGHYDNOQNGIKRKEFLESSPNTQWELRAFMA 120
 QY 316 VMEHSLADRIDDDILKVIYDSNMHHGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 375
 DB 121 VMEHSLADRIDDDILKVIYDSNMHHGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 180
 QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPQTTIOVD-----GSEK 426
 DB 181 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPQTTIKDELHEVGVIAGAEK 240
 QY 427 KIYSIKDFESENKRRGALGNLKNKSYNKNDELSHFATTCSDKSRPLNDVSOXTTQ 486
 DB 241 KIYSIKDFESENKRRGALGNLKNKSYNKNDELSHFATTCSDKSRPLNDVSOXTTQ 300
 QY 487 SDITSRNSAIEALNRFIOKYSVMORLLDDT 518
 DB 301 SDITSRNSAIEALNRFIOKYSVMORLLDDT 332

RP SEQUENCE FROM N.A.
 RC STRAIN-NCIC;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 RT "Passive immunity to infection with *Yersinia* spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96797; CAA65589.1; -
 FT NON_TER 334 334
 SO SEQUENCE 334 AA; 38227 MW; F23PDED3 CRC32;

Query Match 58.1%; Score 1549.5; DB 2; Length 334;
 Best Local Similarity 92.8%; Pred. No. 1.le-74;
 Matches 308; Conservative 9; Mismatches 6; Indels 9; Gaps 1;

QY 196 MIRAYEQNPOHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 255
 DB 1 MIRAYEQNPOHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 60
 QY 256 NRVTDDIELLKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLESSPNTQWELRAFMA 315
 DB 61 ERVTDDIELLKKILAYFLPEDAILKGGHYDNOLONGIKRVEFLESSPNTQWELRAFMA 120
 QY 316 VMHFSLTADRIDDDILKIVYDSNMHGDARSKRLRELAELTAEIKIYVIOAEINKHLSS 375
 DB 121 VMHFSLTADRIDDDILKIVYDSNMHGDARSKRLRELAELTAEIKIYVIOAEINKHLSS 180
 QY 376 SGTINIHDKSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDS-----GSEK 426
 DB 181 SGTINIHDKSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDS-----GSEK 240
 QY 427 KIVSKDGLGSEKRTGALGNKNSYKNNKNNELSHFATTCSDKSRPLNDVYSOKTQOL 486
 DB 241 QIVSKNFESEKRTGALGNKNSYKNNKNNELSHFATTCSDKSRPLNDVYSOKTQOL 300
 QY 487 SDITSRFNSAIEALNRFIOKYDSVMQRLDDT 518
 DB 301 SDITSRFNSAIEALNRFIOKYDSVMQRLDDT 332

RESULT 5
 056880 PRELIMINARY: PRT: 325 AA.

AC 056880;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
 DE V ANTIGEN (FRAGMENT).
 OS *Yersinia enterocolitica*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 NC *Yersinia*.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-8081;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 RT "Passive immunity to infection with *Yersinia* spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96798; CAA65590.1; -
 FT NON_TER 325 325
 SO SEQUENCE 325 AA; 37260 MW; DEE88446 CRC32;

Query Match 57.7%; Score 1537; DB 2; Length 325;
 Best Local Similarity 94.4%; Pred. No. 4.8e-74;
 Matches 305; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 196 MIRAYEQNPOHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 255
 DB 1 MIRAYEQNPOHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 60

QY 256 NRVTDDIELLKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLESSPNTQWELRAFMA 315
 DB 61 ERVTDDIELLKKILAYFLPEDAILKGGHYDNOLONGIKRVEFLESSPNTQWELRAFMA 120
 QY 316 VMHFSLTADRIDDDILKIVYDSNMHGDARSKRLRELAELTAEIKIYVIOAEINKHLSS 375
 DB 121 VMHFSLTADRIDDDILKIVYDSNMHGDARSKRLRELAELTAEIKIYVIOAEINKHLSS 180
 QY 376 SGTINIHDKSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDSSEKKIYSINDEL 435
 DB 181 SGTINIHDKSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDSSEKKIYSINDEL 240
 QY 436 GSEKRTGALGNKNSYKNNKNNELSHFATTCSDKSRPLNDVYSOKTQOLSDITSRENS 495
 DB 241 ESEKRTGALGNKNSYKNNKNNELSHFATTCSDKSRPLNDVYSOKTQOLSDITSRENS 300
 QY 496 AIEALNRFIOKYDSVMQRLDDT 518
 DB 301 AIEALNRFIOKYDSVMQRLDDT 323

RESULT 6
 056882 PRELIMINARY: PRT: 324 AA.

AC 056882;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
 DE V ANTIGEN.
 OS *Yersinia enterocolitica*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 NC *Yersinia*.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y-96-P;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 RT "Passive immunity to infection with *Yersinia* spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96801; CAA65593.1; -
 SO SEQUENCE 324 AA; 37162 MW; 56842D0A CRC32;

Query Match 57.5%; Score 1532; DB 2; Length 324;
 Best Local Similarity 93.8%; Pred. No. 8.8e-74;
 Matches 303; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 196 MIRAYEQNPOHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 255
 DB 1 MIRAYEQNPOHFIEDLEKVRVQOLTGHSVLEELVOLVKDNIDISIKYDPRKDSVEFA 60
 QY 256 NRVTDDIELLKKILAYFLPEDTILKGGHYDNOLONGIKRVEFLESSPNTQWELRAFMA 315
 DB 61 DRVTDDIELLKKILAYFLPEDAILKGGHYDNOLONGIKRVEFLESSPNTQWELRAFMA 120
 QY 316 VMHFSLTADRIDDDILKIVYDSNMHGDARSKRLRELAELTAEIKIYVIOAEINKHLSS 375
 DB 121 VMHFSLTADRIDDDILKIVYDSNMHGDARSKRLRELAELTAEIKIYVIOAEINKHLSS 180
 QY 376 SGTINIHDKSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDSSEKKIYSINDEL 435
 DB 181 SGTINIHDKSINLMDKNLYGTDEEIFRASAERYKILEKMPOTTIOVDSSEKKIYSINDEL 240
 QY 436 GSEKRTGALGNKNSYKNNKNNELSHFATTCSDKSRPLNDVYSOKTQOLSDITSRENS 495
 DB 241 ESEKRTGALGNKNSYKNNKNNELSHFATTCSDKSRPLNDVYSOKTQOLSDITSRENS 300
 QY 496 AIEALNRFIOKYDSVMQRLDDT 518
 DB 301 AIEALNRFIOKYDSVMQRLDDT 323

RESULT 7
 056846 PRELIMINARY: PRT: 324 AA.
 AC 056846:
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, last annotation update)
 DE V ANTIGEN.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y-108-P;
 RX MEDLINE: 97162308.
 RA ROGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96799; CAA65591.1;
 SQ SEQUENCE 324 AA; 37194 MW; 5BD357FF CRC32;

Query Match 57.4%; Score 1531; DB 2; Length 324;
 Best Local Similarity 94.1%; Pred. No. 1e-73;
 Matches 304; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 196 MIRAEONQPHIEDLEKRVQLTGHSSVLEELVQVKKNDISIKYPRKSEVFA 255
 |||||||
 DB 1 MIRAEONQPHIEDLEKRVQLTGHSSVLEELVQVKKNDISIKYPRKSEVFA 60
 |||||||
 QY 256 NRVITDDIELKKIAYFLPEDTIILKGHYDNOLONGIKRVEFLSSPNTOMELRAVFA 315
 |||||||
 DB 61 NRVITDDIELKKIAYFLPEDTIILKGHYDNOLONGIKRVEFLSSPNTOMELRAVFA 120
 |||||||
 QY 316 VVHFSLTADRIDDLIKYIVDSMNHNGDARSKLRELAELTAELKIYVIOAEINKHLS 375
 |||||||
 DB 121 VVHFSLTADRIDDLIKYIVDSMNHNGDARSKLRELAELTAELKIYVIOAEINKHLS 180
 |||||||
 QY 376 SGTINIHDKSINLMKKNLYGYDEIEFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFL 435
 |||||||
 DB 181 SGTINIHDKSINLMKKNLYGYDEIEFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFL 240
 |||||||
 QY 436 GSENRRTGALGNLKNYSYKNDNLSHFATCSDKSRPLNDVSOQTOLSDITSRFS 495
 |||||||
 DB 241 ESENRRTGALGNLKNYSYKNDNLSHFATCSDKSRPLNDVSOQTOLSDITSRFS 300
 |||||||
 QY 496 AIEALNRFIOKYDSVMORLLDDT 518
 |||||||
 DB 301 AIEALNRFIOKYDSVMORLLDDT 323
 |||||||

RESULT 8
 056879 PRELIMINARY: PRT: 324 AA.
 AC 056879:
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, last annotation update)
 DE V ANTIGEN.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97162308.
 RA ROGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
 "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96800; CAA65592.1;
 SQ SEQUENCE 324 AA; 37195 MW; EICB300F CRC32;

Query Match 57.3%; Score 1527; DB 2; Length 324;
 Best Local Similarity 93.8%; Pred. No. 1.6e-73;
 Matches 303; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 196 MIRAEONQPHIEDLEKRVQLTGHSSVLEELVQVKKNDISIKYPRKSEVFA 255
 |||||||
 DB 1 MIRAEONQPHIEDLEKRVQLTGHSSVLEELVQVKKNDISIKYPRKSEVFA 60
 |||||||
 QY 256 NRVITDDIELKKIAYFLPEDTIILKGHYDNOLONGIKRVEFLSSPNTOMELRAVFA 315
 |||||||
 DB 61 NRVITDDIELKKIAYFLPEDTIILKGHYDNOLONGIKRVEFLSSPNTOMELRAVFA 120
 |||||||
 QY 316 VVHFSLTADRIDDLIKYIVDSMNHNGDARSKLRELAELTAELKIYVIOAEINKHLS 375
 |||||||
 DB 121 VVHFSLTADRIDDLIKYIVDSMNHNGDARSKLRELAELTAELKIYVIOAEINKHLS 180
 |||||||
 QY 376 SGTINIHDKSINLMKKNLYGYDEIEFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFL 435
 |||||||
 DB 181 SGTINIHDKSINLMKKNLYGYDEIEFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFL 240
 |||||||
 QY 436 GSENRRTGALGNLKNYSYKNDNLSHFATCSDKSRPLNDVSOQTOLSDITSRFS 495
 |||||||
 DB 241 ESENRRTGALGNLKNYSYKNDNLSHFATCSDKSRPLNDVSOQTOLSDITSRFS 300
 |||||||
 QY 496 AIEALNRFIOKYDSVMORLLDDT 518
 |||||||
 DB 301 AIEALNRFIOKYDSVMORLLDDT 323
 |||||||

RESULT 9
 030527 PRELIMINARY: PRT: 294 AA.
 AC 030527:
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, last annotation update)
 DE PCR.
 GN Pseudomonas aeruginosa.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
 CC Pseudomonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-388;
 RX MEDLINE: 98037517.
 RA YAHK T.L., MENDE-MUELLER L.M., FRIESE M.B., FRANK D.W.;
 "Identification of type III secreted products of the Pseudomonas
 aeruginosa exoenzyme S regulon";
 RL J. Bacteriol. 179:7165-7168(1997).
 DR EMBL: AF010149; AAC45935.1;
 SQ SEQUENCE 294 AA; 32283 MW; 0FAD2082 CRC32;

Query Match 18.4%; Score 490; DB 2; Length 294;
 Best Local Similarity 39.7%; Pred. No. 5.7e-19;
 Matches 117; Conservative 54; Mismatches 92; Indels 32; Gaps 7;

QY 224 SVLEELVQVKKNDISIKYPRKSEVFAANRVITDDIELKKIAYFLPEDTIILKG 283
 |||||||
 DB 26 SVLEELVQVKKNDISIKYPRKSEVFAANRVITDDIELKKIAYFLPEDTIILKG 72
 |||||||
 QY 284 HYDNOLONGIKRVEFLSSPNTOMELRAVFAVHFSLTADRIDDLIKYIVDSMNH 341
 |||||||
 DB 73 HYDNOLONGIKRVEFLSSPNTOMELRAVFAVHFSLTADRIDDLIKYIVDSMNH 124
 |||||||
 QY 342 GDARSKLRELAELTAELKIYVIOAEINKHLSGGTINIHDKSINLMKKNLYGY-1DEE 400
 |||||||
 DB 125 GDARSKLRELAELTAELKIYVIOAEINKHLSGGTINIHDKSINLMKKNLYGY-1DEE 184
 |||||||
 QY 401 IFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFLGSSENRRTGALGNLKNYSYKND 460
 |||||||
 DB 185 IFKASAEYKILEKMPQTTIOVDSSEKRIYSIKDFLGSSENRRTGALGNLKNYSYKND 236
 |||||||

Db 1235 CYGFINSGEIRAGVENIKN-----HF-DVCIKKNTLDDIYKYE----- 1277
 QY 487 SDITSRNSAIEALNRFQKYSVMORLDDTSG 520
 Db 1278 ----NEIHKRIDSLKSIESTFSDIEKMLNDKVSg 1307

RESULT 12

ID 090763 PRELIMINARY; PRT; 785 AA.
 AC 090763;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
 DE CHICKEN CADHERIN-7.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
 OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WHITE LEGHORN; TISSUE-BRAIN;
 RA MEDLINE; 95309115.
 RX NAKAGAWA S., TAKEICHI M.;
 RT "Neural crest cell-cell adhesion controlled by sequential and
 subpopulation-specific expression of novel cadherins."
 RL Development 121:1321-1332(1995).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL; D42150; BAA07721.1; -.
 DR HSSP; P15116; INCY.
 DR PROSITE; PS00232; CADHERIN; 3.
 DR PFAM; PF00028; cadherin; 5.
 DR PFAM; PF01049; Cadherin_C-term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.
 KW SEQUENCE 785 AA; 87171 MW; F27A6881 CRC32;

Query Match 5.5%; Score 147; DB 13; Length 785;
 Best Local Similarity 23.0%; Pred. No. 2.6; Mismatches 26;

Matches 115; Conservative 71; Mismatches 201; Indels 114; Gaps 26;
 QY 12 HSSGHIDDDDKHKKI-SSVIAIALFETIATANAADLTAS-----TTATATLVEPARIT 64
 Db 72 HSD -VDKGGSGIKYILISGASISF--IIDENGTGDIHAIRKLDREQATYTLRAQADR 127
 QY 65 LTYKEAP-----TTMDNGNIDTELLVGLTGK--TGTSTSVNFTDAAGPMYL 116
 Db 128 LTRKPYEPESEFYIKIODINDNEPKFLDGYTAGVPEMSPVGTSVQVATD-ADPTY- 185
 QY 117 TFSOGNNHQTXYIGKSRDFDISPK-----NENLVGDVYVATSGDFVR 168
 Db 186 -----GNSARVYSL-QQPYFSEVPKGIKTLALPNDRKAKQYLLVIAKMD-- 235
 QY 169 SIGSGKGLAGKYTDAVTVVTSNOEFMIRAYONPOHFIEDLEKRVDELJGHSSYLE 228
 Db 236 -VGONG--LSG--TTSVYTLTD-----VNDNPRFRASYQVNP-----SLPLA 278
 QY 229 ELVOLYKKNIDISIKYPRKDEVFANRYITDIELKILAYFLPEDTILKGHYDQ 288
 Db 279 SVARAKADADV-----PVAEMEY--KIVDGDGLGVFI-----SVDKD 317
 QY 289 LONGIRKVEFLESSPRTQELRAFAVMA-----FSLTADRIDDILKIVYDSMHHGDA 344
 Db 318 TEGGITTKEIDFEAKTSTLRIEAMNVDPRFLSLGFSMTYVKIIVEDV----- 371
 QY 345 RSKRLRELAELTALIKYVIOAEINKHSSSGTINIHDSIN-----LMDNLYGYTD 398
 Db 372 -----DEPVFTS--RLYSVVEAAKVGTTIGTVAHDDASNPYRISIDRN--TD 420
 QY 399 -EEIFRASAERYKLEKMPOTTIOVDSEKKIVSIRKDFLGSNNRTGALNKLKSYSNKD 457
 Db 421 LERYFNIDANSYI-----TTAKSLDRETNVAINITYLAMESQNPAPDIGHGYAIIITLDI 475

QY 458 NNELSHFA-----TTCGDSKRP 474
 Db 476 NDNAPEFAMEYETTVCENAP 496

RESULT 13

ID 094488 PRELIMINARY; PRT; 1115 AA.
 AC 094488;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
 DE PUTATIVE COILED-COIL PROTEIN.
 GN SPC417.07C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972H-;
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKAERT G.;
 RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; At035076; CAA22653.1; -.
 KW SEQUENCE 1115 AA; 128468 MW; 6E182EE3 CRC32;

Query Match 5.4%; Score 144; DB 3; Length 1115;
 Best Local Similarity 18.6%; Pred. No. 6;

Matches 123; Conservative 100; Mismatches 180; Indels 258; Gaps 29;

QY 16 HIDDDE-----H-----MKISSVIAIALFETIATANAADLTASTT 52
 Db 182 HHDEDDDLQTPKEERNYLSQDAPPTNALSKISDIL-----IPASAKMLDORKN 234
 QY 53 ATATLVEPAR--TLTYKKGAPITMDN----- 78
 Db 235 ALAKEPEESOPGSSLTLLKQDA--NVIDNLRKEVFGKLKCYLYDQLNFHQEVDINK 292
 QY 79 GNIDTELLVGLT--TLGKYKGTSTSVNFTDAAGPMYITFTSOGNNHQTXYIGKD 136
 Db 293 QNIDKTLIMELORAVAGE-----KKSIGLE 319
 QY 137 SR-----DFDIS--PKVNGENLVGD-----VIALTSGDF--FVRSIGSGKGLAAGY 182
 Db 320 SRIKPDQSNLSTPPSPAPSNLITLQSRYSQALSELETTKRAFALAKREKSKTNSVGA 379
 QY 183 TDVATVTSNOEFMIRAYONPOHFIEDLEKRVDELJGHSSYLELVOLYKKNIDIS 242
 Db 380 NEDRN-LSN--MDNERREKALQELSLRVO-----LSKRYPM----- 417
 QY 243 IKYDPKDESEVFANRYITDIELKILAYFLPEDTILKGHYDNLONGIRKVEF-- 299
 Db 418 -----PAKNTERVETLQSNELLRDIS--KONELLARKOENRDL--VQVVELITVA 468
 QY 300 -----LESSPNTQWE--LRAFAVMAHSLTADR-----IDDLIKIYVDSMN 339
 Db 469 LNSGKNAIYEAESSKNELMDSMWYSRMKTQOSILFRLYQLODIEDYENKILRMQD 528
 QY 340 HHGDAASKURELAELTAL-----KIYVIOAEINKHLSSSG 377
 Db 529 QWRVEDVQLOEYVEITLOLQDTEKVLSSKSESDYEYVGKLRTEAREIEKF--EK 585
 QY 378 TINIHKSNIMDKNLYGTDEEIFKASAEY----- 408
 Db 586 TITENESSISLREVEKLT-D-IYOLSRVYDKCHEPDELQKRLQTLTEENNAKEDST 644
 QY 409 ----KILEKMPOTTIOVDSEKKIVSIRKDFLGSNNRTGALNKLKSYSNKD 464
 Db 645 SKTSNLEQLKTEAVDLSRKE-----NENKQVIA--LKESELVSKNDKL-- 690
 QY 465 ATTCSPKSPRLNDLVYQKTLQSDITS-----RNSAIEALNRFQKYSV 510

Wed Aug 23 11:46:40 2000

us-08-699-716a-2.rspt

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 15:11:47 ; Search time 1417.37 Seconds
(without alignments)
1971.845 Million cell updates/sec

Title: US-08-699-716a-1
Perfect score: 1566
Sequence: 1 ATGGCCATCATCATCATCA.....ATGACACGTCGTGTAATCA 1566

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 972840 seqs, 892348106 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl:*
1: gb_da1:*
2: gb_da2:*
3: gb_cm:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_p11:*
8: gb_p12:*
9: gb_p13:*
10: gb_p14:*
11: gb_p15:*
12: gb_p16:*
13: gb_p17:*
14: gb_p18:*
15: gb_p19:*
16: gb_p20:*
17: gb_p21:*
18: gb_p22:*
19: gb_p23:*
20: gb_p24:*
21: gb_p25:*
22: gb_p26:*
23: gb_p27:*
24: gb_p28:*
25: gb_p29:*
26: gb_p30:*
27: gb_p31:*
28: gb_p32:*
29: gb_p33:*
30: gb_p34:*
31: gb_p35:*
32: gb_p36:*
33: gb_p37:*
34: gb_p38:*
35: gb_p39:*
36: gb_p40:*
37: gb_p41:*
38: gb_p42:*
39: gb_p43:*
40: gb_p44:*
41: gb_p45:*
42: gb_p46:*
43: gb_p47:*

44: gb_p48:*
45: gb_p49:*
46: gb_p50:*
47: gb_p51:*
48: gb_p52:*
49: gb_p53:*
50: gb_p54:*
51: gb_p55:*
52: gb_p56:*
53: gb_p57:*
54: gb_p58:*
55: gb_p59:*
56: gb_p60:*
57: gb_p61:*
58: gb_p62:*
59: gb_p63:*
60: gb_p64:*
61: gb_p65:*
62: gb_p66:*
63: gb_p67:*
64: gb_p68:*
65: gb_p69:*
66: gb_p70:*
67: gb_p71:*
68: gb_p72:*
69: gb_p73:*
70: gb_p74:*
71: gb_p75:*
72: gb_p76:*
73: gb_p77:*
74: gb_p78:*
75: gb_p79:*
76: gb_p80:*
77: gb_p81:*
78: gb_p82:*
79: gb_p83:*
80: gb_p84:*
81: gb_p85:*
82: gb_p86:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	930	59.4	2100	1 YPDCR
2	930	59.4	70305	2 YPDC1
3	930	59.4	70504	2 AF053946
4	930	59.4	70559	2 AF074612
5	927	59.2	1014	5 A46411
6	927	59.2	1014	5 A56793
7	927	59.2	1462	5 A56808
8	927	59.2	1530	5 A56814
9	923	58.9	1014	5 A56795
10	923	58.6	1014	5 A56795
11	918	58.6	981	2 AF167309
12	777	49.6	981	2 AF167310
13	513	32.8	547	5 A56812
14	512	32.7	5383	1 YPCAF
15	512	32.7	96210	2 YPPMT1
16	512	32.7	100984	2 AF053947
17	512	32.7	100990	2 AF074611
18	438	28.0	981	1 YPPTVANT
19	392	25.0	2201	1 YPDCRGVHP
20	259	16.5	1002	1 YP314VANT
21	215	13.7	1002	1 YP314VANT
22	196	12.5	975	1 YP108VANT
23	196	12.5	975	1 YP527VANT
24	189	12.1	975	1 YP966VANT

C	25	189	12.1	69673	2	AF102990	AF102990	Yersinia	98	30	1.9	332	13	G09256	G09256	human	STS	C
C	26	146	9.3	975	1	YEB08VANT	X96798	Y. enterocol	99	30	1.9	334	13	G10329	G10329	human	STS	C
C	27	74	4.7	660	14	AF081364	AF081364	Synthetic	100	30	1.9	335	13	G09261	G09261	human	STS	C
C	28	53	3.4	240	1	S38727	S38727	lcrGVH oper	101	30	1.9	349	13	G09257	G09257	human	STS	C
C	29	53	3.4	521	10	S75264	S75264	Wtl-Wtlms'	102	30	1.9	360	34	AF169216	AF169216	human	STS	C
C	30	53	3.4	1402	5	AR000030	AR000030	Sequence	103	30	1.9	366	13	AF169216	AF169216	human	STS	C
C	31	35	2.2	12800	39	AC005192	AC005192	Homo sapi	104	30	1.9	377	13	G09253	G09253	human	STS	C
C	32	34	2.2	63654	49	ATAC01698	AC01698	Arabidops	105	30	1.9	395	13	G08663	G08663	human	STS	C
C	33	34	2.2	110804	49	ATAC016829	AC016829	Arabidops	106	30	1.9	397	13	G09248	G09248	human	STS	C
C	34	33	2.1	67213	53	AC024522	AC024522	Homo sapi	107	30	1.9	403	13	G09249	G09249	human	STS	C
C	35	33	2.1	151212	32	AL136985	AL136985	Homo sapi	108	30	1.9	444	13	H0006615	H0006615	human	STS	C
C	36	33	2.1	166765	39	AC009721	AC009721	Homo sapi	109	30	1.9	559	13	G15746	G15746	human	STS	C
C	37	32	2.0	314	13	G08591	G08591	human	110	30	1.9	634	13	H0006629	H0006629	human	STS	C
C	38	32	2.0	521	13	H0007996	L30172	Human	111	30	1.9	964	12	AF074963	AF074963	human	STS	C
C	39	32	2.0	5017	39	COL1A1G12	AF010190	Homo sapi	112	30	1.9	1099	34	AF030304	AF030304	human	STS	C
C	40	32	2.0	30676	4	AF089743	AF089743	Morone sa	113	30	1.9	1644	3	DDU65018	DDU65018	human	STS	C
C	41	32	2.0	68102	51	AC023751	AC023751	Homo sapi	114	30	1.9	1839	7	SCU12825	SCU12825	human	STS	C
C	42	32	2.0	83263	80	AC063954	AC063954	Homo sapi	115	30	1.9	2204	39	AF091095	AF091095	human	STS	C
C	43	32	2.0	110535	39	AC005091	AC005091	Homo sapi	116	30	1.9	2836	7	SCYK1028W	SCYK1028W	human	STS	C
C	44	32	2.0	127425	39	AC004976	AC004976	Homo sapi	117	30	1.9	4381	12	RNCAMP9A	RNCAMP9A	human	STS	C
C	45	32	2.0	145831	39	H00A001548	AF001548	Human	118	30	1.9	4596	6	SCYOL138C	SCYOL138C	human	STS	C
C	46	32	2.0	157258	50	AC008279	AC008279	Homo sapi	119	30	1.9	4779	9	AB023203	AB023203	human	STS	C
C	47	32	2.0	166633	54	AC012568	AC012568	Homo sapi	120	30	1.9	7176	33	PFACDST	PFACDST	human	STS	C
C	48	31	2.0	255	13	PCU64650	U64650	Paracharter	121	30	1.9	7407	10	D495554	D495554	human	STS	C
C	49	31	2.0	256	13	G08117	G08117	human	122	30	1.9	8152	54	AC023164	AC023164	human	STS	C
C	50	31	2.0	451	13	G09413	G09413	human	123	30	1.9	8920	5	AR065855	AR065855	Sequence		
C	51	31	2.0	588	13	G09471	G09471	human	124	30	1.9	12805	7	SC128XV	SC128XV	human	STS	C
C	52	31	2.0	1867	33	DDIPK1	G09471	human	125	30	1.9	15372	33	DMBR27C19	DMBR27C19	human	STS	C
C	53	31	2.0	38182	43	AC019948	AC019948	Drosophila	126	30	1.9	16078	43	AC019969	AC019969	human	STS	C
C	54	31	2.0	47812	11	AC005351	AC005351	Homo sapi	127	30	1.9	16068	42	AC013095	AC013095	human	STS	C
C	55	31	2.0	63402	9	AP000260	AP000260	Homo sapi	128	30	1.9	20090	11	HSBA59J12	HSBA59J12	human	STS	C
C	56	31	2.0	69696	51	AC023030	AC023030	Homo sapi	129	30	1.9	25759	12	RNMHG	RNMHG	human	STS	C
C	57	31	2.0	90223	39	AC000105	AC000105	Homo sapi	130	30	1.9	28142	10	HS158B6	HS158B6	human	STS	C
C	58	31	2.0	93163	11	HSAC000121	AC000121	Human	131	30	1.9	29632	131	HS215F16	HS215F16	human	STS	C
C	59	31	2.0	100000	9	AP000099	AP000099	Homo sapi	132	30	1.9	31071	51	AC023723	AC023723	human	STS	C
C	60	31	2.0	105726	12	AF037352	AF037352	Homo sapi	133	30	1.9	38461	10	HSU165H7	HSU165H7	human	STS	C
C	61	31	2.0	135160	54	AC008449	AC008449	Homo sapi	134	30	1.9	39739	42	AC015123	AC015123	human	STS	C
C	62	31	2.0	137713	74	AC027683	AC027683	Homo sapi	135	30	1.9	39960	10	HSU162C4	HSU162C4	human	STS	C
C	63	31	2.0	144857	9	AP000036	AP000036	Homo sapi	136	30	1.9	41572	39	AC006132	AC006132	human	STS	C
C	64	31	2.0	144857	9	AP000036	AP000036	Homo sapi	137	30	1.9	45296	39	AC007746	AC007746	human	STS	C
C	65	31	2.0	145824	77	AC018441	AC018441	Homo sapi	138	30	1.9	48713	78	AC060788	AC060788	human	STS	C
C	66	31	2.0	146721	55	AC010303	AC010303	Homo sapi	139	30	1.9	49648	7	AP000607	AP000607	human	STS	C
C	67	31	2.0	155867	11	HS404H4	AL031661	Human	140	30	1.9	61419	42	AC014418	AC014418	human	STS	C
C	68	31	2.0	159500	12	AC005742	AC005742	Mus muscu	141	30	1.9	65443	42	AC014845	AC014845	human	STS	C
C	69	31	2.0	163352	39	AC009225	AC009225	Homo sapi	142	30	1.9	68141	70	AC027361	AC027361	human	STS	C
C	70	31	2.0	169236	10	HSJ1123D4	AC046197	Human	143	30	1.9	68724	76	AC046197	AC046197	human	STS	C
C	71	31	2.0	169765	39	AC006075	AC006075	Homo sapi	144	30	1.9	69119	75	AC048366	AC048366	human	STS	C
C	72	31	2.0	171624	40	AL160283	AL160283	Homo sapi	145	30	1.9	69847	50	AC004845	AC004845	human	STS	C
C	73	31	2.0	172367	59	AC024457	AC024457	Homo sapi	146	30	1.9	71894	60	AC025810	AC025810	human	STS	C
C	74	31	2.0	175470	78	AC016903	AC016903	Homo sapi	147	30	1.9	73844	10	HSBK109D1	HSBK109D1	human	STS	C
C	75	31	2.0	178814	67	AC022927	AC022927	Homo sapi	148	30	1.9	74546	42	AC0138919	AC0138919	human	STS	C
C	76	31	2.0	179281	57	AC019328	AC019328	Homo sapi	149	30	1.9	76526	32	AC0138912	AC0138912	human	STS	C
C	77	31	2.0	179583	52	AC020628	AC020628	Homo sapi	150	30	1.9	77227	55	AC025099	AC025099	human	STS	C
C	78	31	2.0	181795	11	AC003685	AC003685	Homo sapi	151	30	1.9	77401	7	AB026657	AB026657	human	STS	C
C	79	31	2.0	189765	10	HS353E16	AL031591	Human	152	30	1.9	78174	34	AC004280	AC004280	human	STS	C
C	80	31	2.0	190216	59	AC024731	AC024731	Homo sapi	153	30	1.9	79279	41	AC011441	AC011441	human	STS	C
C	81	31	2.0	193822	76	AC021132	AC021132	Homo sapi	154	30	1.9	80209	11	AC000370	AC000370	human	STS	C
C	82	31	2.0	196382	51	AC020886	AC020886	Mus muscu	155	30	1.9	81494	7	AB019236	AB019236	human	STS	C
C	83	31	2.0	199723	42	AC016756	AC016756	Homo sapi	156	30	1.9	85633	42	AC014423	AC014423	human	STS	C
C	84	31	2.0	200685	39	AC004032	AC004032	Homo sapi	157	30	1.9	87077	39	H00AC04097	H00AC04097	human	STS	C
C	85	31	2.0	205228	32	AL137219	AL137219	Homo sapi	158	30	1.9	87459	7	AB015477	AB015477	human	STS	C
C	86	31	2.0	210616	56	AC024113	AC024113	Mus muscu	159	30	1.9	89479	49	AC006932	AC006932	Genomic s		
C	87	31	2.0	210882	55	AC020910	AC020910	Homo sapi	160	30	1.9	91470	49	AC011438	AC011438	human	STS	C
C	88	31	2.0	213011	53	AC018664	AC018664	Homo sapi	161	30	1.9	92073	72	AC008968	AC008968	human	STS	C
C	89	31	2.0	223998	40	AL139338	AL139338	Homo sapi	162	30	1.9	92657	8	ATT15N24	ATT15N24	human	STS	C
C	90	31	2.0	232275	79	AC023811	AC023811	Mus muscu	163	30	1.9	92911	43	AC020218	AC020218	human	STS	C
C	91	31	2.0	304060	34	AE003447	AE003447	Drosophila	164	30	1.9	92944	32	HSU242F8	HSU242F8	human	STS	C
C	92	31	2.0	318377	12	RNEDNRB7	AC010955	Homo sapi	165	30	1.9	100209	10	HS33L1	HS33L1	human	STS	C
C	93	30	1.9	115	12	AF027614	U93687	Rattus norv	166	30	1.9	102064	11	AC002558	AC002558	human	STS	C
C	94	30	1.9	189	8	AF023249	AF023249	Fraxinus	167	30	1.9	102077	8	FL1A17	FL1A17	human	STS	C
C	95	30	1.9	205	33	PLATRE	L10836	Polistes an	168	30	1.9	107527	39	AC005018	AC005018	human	STS	C
C	96	30	1.9	326	13	H0006005B	L30178	Human	169	30	1.9	109891	10	HS112K5	HS112K5	human	STS	C
C	97	30	1.9	326	13	H0006005B	L30178	Human	170	30	1.9	110469	8	AC006258	AC006258	human	STS	C

171	30	1.9	111222	49	AC007764	Genomic s
172	30	1.9	111368	9	AP000025	Hom sapi
173	30	1.9	117767	40	AL160262	Hom sapi
174	30	1.9	119118	10	HSR97CC1	Hom sapi
175	30	1.9	120088	39	AC004865	Hom sapi
176	30	1.9	120937	55	AC010239	Hom sapi
177	30	1.9	121695	32	AL137127	Hom sapi
178	30	1.9	121990	51	AC022314	Arabidops
179	30	1.9	122264	32	HSAL31498	Hom sapi
180	30	1.9	123762	42	AF207954	Hom sapi
181	30	1.9	128293	67	AC016564	Hom sapi
182	30	1.9	128577	67	AC025451	Hom sapi
183	30	1.9	129661	52	AC018543	Hom sapi
184	30	1.9	131007	10	HS998C11	Human DNA
185	30	1.9	131819	11	HS529N6	Human DNA
186	30	1.9	131899	43	AC008258	Drosophila
187	30	1.9	132723	61	AC009848	Drosophila
188	30	1.9	134019	60	AC021361	Hom sapi
189	30	1.9	134593	67	AC008135	Drosophila
190	30	1.9	137938	60	AC011653	Hom sapi
191	30	1.9	138828	43	AC016468	Hom sapi
192	30	1.9	139498	73	AC025297	Hom sapi
193	30	1.9	141762	10	HS198C21	Human DNA
194	30	1.9	145573	32	AL138904	Hom sapi
195	30	1.9	147157	10	HS545L17	Human DNA
196	30	1.9	149441	55	AC010432	Hom sapi
197	30	1.9	149560	43	AC016050	Hom sapi
198	30	1.9	149560	43	AC016050	Hom sapi
199	30	1.9	150204	41	AF186190	Hom sapi
200	30	1.9	150589	78	AC021360	Hom sapi
201	30	1.9	152192	40	AL158814	Hom sapi
202	30	1.9	152448	71	AC021391	Hom sapi
203	30	1.9	153987	80	AC066616	Hom sapi
204	30	1.9	153987	80	AC066616	Hom sapi
205	30	1.9	153987	80	AC066616	Hom sapi
206	30	1.9	153987	80	AC066616	Hom sapi
207	30	1.9	157642	54	AC022865	Hom sapi
208	30	1.9	158184	53	AC024158	Hom sapi
209	30	1.9	158499	57	AC011642	Hom sapi
210	30	1.9	162698	72	AC011070	Hom sapi
211	30	1.9	163266	73	AC008696	Hom sapi
212	30	1.9	163277	76	AC010095	Hom sapi
213	30	1.9	164132	60	AC019284	Hom sapi
214	30	1.9	164399	33	PFMAL3P6	Plasmodiu
215	30	1.9	164762	74	AC016177	Hom sapi
216	30	1.9	165010	11	HSDB39B4	Human DNA
217	30	1.9	165047	55	AC025284	Hom sapi
218	30	1.9	165197	11	HSAC002070	Human BAC
219	30	1.9	165429	40	AL157372	Hom sapi
220	30	1.9	165921	53	AC019354	Hom sapi
221	30	1.9	165947	68	AC015897	Hom sapi
222	30	1.9	166016	40	AL159999	Hom sapi
223	30	1.9	166629	80	AC064847	Hom sapi
224	30	1.9	167070	44	AC009969	Hom sapi
225	30	1.9	167241	41	AL153764	Hom sapi
226	30	1.9	167271	53	AC006262	Hom sapi
227	30	1.9	167397	70	AC007525	Hom sapi
228	30	1.9	168335	55	AC015662	Hom sapi
229	30	1.9	168835	10	HS652L8	Human DNA
230	30	1.9	169811	55	AC011454	Hom sapi
231	30	1.9	170028	69	AC027271	Hom sapi
232	30	1.9	170209	32	AL138815	Hom sapi
233	30	1.9	170610	61	AC019329	Hom sapi
234	30	1.9	171660	32	AL139283	Hom sapi
235	30	1.9	172339	70	AC007430	Hom sapi
236	30	1.9	172436	70	AC025854	Hom sapi
237	30	1.9	172793	69	AC023659	Hom sapi
238	30	1.9	173126	79	AC005988	Hom sapi
239	30	1.9	173640	79	AC024479	Hom sapi
240	30	1.9	175799	40	AL162426	Hom sapi
241	30	1.9	176326	73	AC025881	Hom sapi
242	30	1.9	176567	43	AF192304	Hom sapi
243	30	1.9	176697	59	AC021170	Hom sapi
244	30	1.9	177292	68	AC010768	Hom sapi
245	30	1.9	180324	67	AC026212	Hom sapi
246	30	1.9	181098	11	AC004216	Hom sapi
247	30	1.9	181191	69	AC022135	Hom sapi
248	30	1.9	181727	67	AC026291	Hom sapi
249	30	1.9	183007	54	AC021540	Hom sapi
250	30	1.9	183372	73	AC009161	Hom sapi
251	30	1.9	183433	60	AC021907	Hom sapi
252	30	1.9	185155	40	AL162414	Hom sapi
253	30	1.9	186206	43	AC009709	Hom sapi
254	30	1.9	187240	42	AC015468	Hom sapi
255	30	1.9	187067	51	AC010066	Hom sapi
256	30	1.9	187956	72	AC021842	Hom sapi
257	30	1.9	188950	77	AC058795	Hom sapi
258	30	1.9	190626	79	AC062036	Hom sapi
259	30	1.9	191479	51	AC023173	Mus muscu
260	30	1.9	191562	40	AL139350	Hom sapi
261	30	1.9	192751	40	AL158159	Hom sapi
262	30	1.9	194275	67	AC025004	Hom sapi
263	30	1.9	194678	71	AC026225	Hom sapi
264	30	1.9	197030	41	AC007914	Hom sapi
265	30	1.9	198330	58	AE003760	Hom sapi
266	30	1.9	198320	8	ATCHRIV65	Hom sapi
267	30	1.9	198549	41	AC009303	Hom sapi
268	30	1.9	198703	59	AC017062	Hom sapi
269	30	1.9	202672	55	AC008827	Hom sapi
270	30	1.9	203348	41	HSBA539A6	Hom sapi
271	30	1.9	203418	75	AC026011	Hom sapi
272	30	1.9	208412	43	AC015557	Hom sapi
273	30	1.9	209645	54	AC015575	Hom sapi
274	30	1.9	209826	74	AC021130	Hom sapi
275	30	1.9	212176	39	AC007543	Hom sapi
276	30	1.9	220210	40	AL136360	Hom sapi
277	30	1.9	222302	32	CNSO1DW3	Hom sapi
278	30	1.9	224840	67	AC008763	Hom sapi
279	30	1.9	224898	79	AC063967	Mus muscu
280	30	1.9	225084	34	AE003687	Mus muscu
281	30	1.9	227106	32	AL139331	Hom sapi
282	30	1.9	227923	44	AC010147	Hom sapi
283	30	1.9	233286	32	AL158079	Hom sapi
284	30	1.9	234112	33	PFMAL4P2	Plasmodiu
285	30	1.9	239285	32	AL157366	Hom sapi
286	30	1.9	248423	42	AC012333	Hom sapi
287	30	1.9	250599	11	AF015262	Hom sapi
288	30	1.9	262082	58	AE003841	Drosophila
289	30	1.9	263712	34	AE003528	Drosophila
290	30	1.9	268002	9	HS229043	Human BAC
291	30	1.9	298873	32	AL158145	Hom sapi
292	30	1.9	299275	34	AE003419	Hom sapi
293	30	1.9	300732	34	AE003450	Hom sapi
294	30	1.9	301550	34	AE003450	Hom sapi
295	29	1.9	301550	34	AE003450	Hom sapi
296	29	1.9	181	13	DDU17368	Human SRS C
297	29	1.9	189	8	G08730	Human SRS C
298	29	1.9	199	13	PF054249	Human SRS C
299	29	1.9	224	13	PF054249	Human SRS C
300	29	1.9	239	13	G08851	Human SRS C
301	29	1.9	250	13	G29268	Human SRS C
302	29	1.9	251	13	PCU64657	Human SRS C
303	29	1.9	269	13	AF193535	Human SRS C
304	29	1.9	271	12	MGU79315	Human SRS C
305	29	1.9	271	12	MGU79315	Human SRS C
306	29	1.9	277	13	MMIGM035	Human SRS C
307	29	1.9	288	13	MMIGM035	Human SRS C
308	29	1.9	298	13	HSB346B5	Human SRS C
309	29	1.9	300	4	MGU79376	Human SRS C
310	29	1.9	311	13	G10330	Human SRS C
311	29	1.9	317	34	AF139021	Human SRS C
312	29	1.9	321	13	G10327	Human SRS C
313	29	1.9	325	13	G08222	Human SRS C
314	29	1.9	342	13	HUM077997	Human SRS C
315	29	1.9	352	13	U0028829	Human SRS C
316	29	1.9	361	13	G09252	Human SRS C

C 317	29	1.9	363	13	G10350	human	STS	C	390	29	1.9	9154	42	AC014305	AC014305 Drosophila	
C 318	29	1.9	367	13	G31855	SPAC1	Notwa	C	391	29	1.9	12029	34	AE001381	AE001381 Plasmodium	
C 319	29	1.9	368	13	G08786	human	STS	C	392	29	1.9	20445	39	AC000023	AC000023 Homo sapi	
C 320	29	1.9	377	13	G27240	human	STS	C	393	29	1.9	21961	33	CER14P4	AL021446 Caenorhab	
C 321	29	1.9	392	34	AF140081	Apis	meil	C	394	29	1.9	22209	12	AF100154	AF100154 Rattus no	
C 322	29	1.9	406	13	G09250	human	STS	C	395	29	1.9	23404	42	AC017160	AC017160 Drosophila	
C 323	29	1.9	410	13	G08916	human	STS	C	396	29	1.9	24197	42	AC015221	AC015221 Drosophila	
C 324	29	1.9	412	34	AF077591	Pennaeus	m	C	397	29	1.9	24942	8	YSCP9819	YSCP9819 Saccharomyc	
C 325	29	1.9	426	34	AF169213	Loligo	pe	C	398	29	1.9	25789	7	SCDWACHV	X91067 S.cerevisia	
C 326	29	1.9	445	13	G41485	Z4192	Zebra	C	399	29	1.9	25937	68	AC026346	AC026346 Homo sapi	
C 327	29	1.9	445	13	G45889	Z4822.1	Zeb	C	400	29	1.9	27470	4	FRDGEWE	X82339 F.rubripes	
C 328	29	1.9	446	13	G45875	Z4740.1	Zeb	C	401	29	1.9	29498	7	AB023029	AB023029 Arabidops	
C 329	29	1.9	453	34	AF169214	Loligo	pe	C	402	29	1.9	29544	33	CERK849	Z82095 Caenorhabdi	
C 330	29	1.9	468	34	AF140078	Apis	meil	C	403	29	1.9	29605	7	AB020747	AB020747 Arabidops	
C 331	29	1.9	500	13	CNS010RK	AL156480	Anopheles	C	404	29	1.9	30578	7	AB017062	AB017062 Arabidops	
C 332	29	1.9	510	13	HDMUT6622	Human	STS	U	405	29	1.9	31205	33	CER106G1	U41014 Caenorhabdi	
C 333	29	1.9	522	12	MMY14111	Mus	musculu	C	406	29	1.9	31527	42	AC012876	AC012876 Drosophila	
C 334	29	1.9	543	34	AF169207	Loligo	pe	C	407	29	1.9	33606	42	AC017861	AC017861 Drosophila	
C 335	29	1.9	558	5	E02043	untransl		C	408	29	1.9	34506	7	SC8337	U73649 Human Chrom	
C 336	29	1.9	612	13	AU028405	Rattus	no	C	409	29	1.9	34864	11	U73649	U73649 Human Chrom	
C 337	29	1.9	625	13	G09478	human	STS	C	410	29	1.9	35576	41	AC007046	AC007046 Homo sapi	
C 338	29	1.9	631	13	G08223	human	STS	C	411	29	1.9	36170	41	AC007657	AC007657 Homo sapi	
C 339	29	1.9	638	13	G39764	Z10215	Zeb	C	412	29	1.9	37353	33	CER1298	Z81120 Caenorhabdi	
C 340	29	1.9	638	13	G39764	Z10215	Zeb	C	413	29	1.9	39387	11	AC004501	AC004501 Homo sapi	
C 341	29	1.9	661	13	G09464	human	STS	C	414	29	1.9	40822	10	HS5F18	Z68694 Human DNA s	
C 342	29	1.9	670	7	R1CSINER12	Rice	retrop	C	415	29	1.9	41917	10	HS5F18	AL035658 Human DNA	
C 343	29	1.9	724	13	G09445	human	STS	C	416	29	1.9	45410	41	AC006187	AC006187 Homo sapi	
C 344	29	1.9	799	13	G40184	Z20576	Zeb	C	417	29	1.9	46934	39	AC010077	AC010077 Homo sapi	
C 345	29	1.9	820	13	G09414	human	STS	C	418	29	1.9	47302	42	AC013346	AC013346 Homo sapi	
C 346	29	1.9	905	8	AF030132	Arabidops		C	419	29	1.9	54368	78	AC062014	AC062014 Homo sapi	
C 347	29	1.9	914	39	AF164668	Homo sapi		C	420	29	1.9	57112	43	AC019976	AC019976 Drosophila	
C 348	29	1.9	926	13	CNS0110E	Onchophiles		C	421	29	1.9	59762	7	AB023032	AB023032 Arabidops	
C 349	29	1.9	1021	81	ORCPV160	Cowpox	viru	C	422	29	1.9	61460	42	AC014853	AC014853 Drosophila	
C 350	29	1.9	1063	81	HAU67261	Helicoverpa		C	423	29	1.9	61756	42	AC017569	AC017569 Drosophila	
C 351	29	1.9	1337	34	AF114166	Cryptospor		C	424	29	1.9	63250	42	AC015767	AC015767 Homo sapi	
C 352	29	1.9	1508	58	S55234	Dictyostell		C	425	29	1.9	64392	78	AC060795	AC060795 Homo sapi	
C 353	29	1.9	1585	11	HSX800565	Dictyostell		C	426	29	1.9	65577	51	AC023689	AC023689 Drosophila	
C 354	29	1.9	1650	12	AF046911	Mus	muscu	C	427	29	1.9	65796	51	AC023570	AC023570 Homo sapi	
C 355	29	1.9	1733	7	ACAG772	Amphidini		C	428	29	1.9	66940	45	AC006242	AC006242 Drosophila	
C 356	29	1.9	1792	33	S76003	S76003	lens-specific	C	429	29	1.9	70447	57	AC025722	AC025722 Caenorhab	
C 357	29	1.9	1792	34	S97760	S-crysta11		C	430	29	1.9	71345	75	AB025613	AB025613 Arabidops	
C 358	29	1.9	1813	12	MUSIGHV2	Mouse	Ig	ge	C	431	29	1.9	71807	7	AB025613	AB025613 Arabidops
C 359	29	1.9	1888	7	ATHUBIOF	Arabidops		C	432	29	1.9	72068	76	AC025785	AC025785 Homo sapi	
C 360	29	1.9	1920	33	OMWCRSBI	M74321	O.sloanei	S	C	433	29	1.9	72456	71	AC027728	AC027728 Homo sapi
C 361	29	1.9	2306	33	DDV18	Al15382	D.alscoiden		C	434	29	1.9	72629	52	AC010174	AC010174 Homo sapi
C 362	29	1.9	2315	11	HSX800321	Homo sapi		C	435	29	1.9	74516	57	AC025641	AC025641 Homo sapi	
C 363	29	1.9	2316	4	ICRIGH	M74041	I.puncatus		C	436	29	1.9	74877	7	AB006705	AB006705 Arabidops
C 364	29	1.9	2359	5	A37615	Sequence	1	C	437	29	1.9	75596	72	AC034122	AC034122 Mus musculu	
C 365	29	1.9	2359	33	CPWALPR	M95743	C.paryum	ge	C	438	29	1.9	77834	43	AC021610	AC021610 Homo sapi
C 366	29	1.9	2359	33	CYDNDNA	M60052	Human	histi	C	439	29	1.9	79469	49	ATAC010852	ATAC010852 Arabidops
C 367	29	1.9	2365	10	HDMHCPB	AF124454	Schizocoe		C	440	29	1.9	79597	39	AF157623	AF157623 Homo sapi
C 368	29	1.9	2392	34	AF124454	Y14756	Mus musculu		C	441	29	1.9	80417	70	AC027562	AC027562 Homo sapi
C 369	29	1.9	2475	12	MMSTOEX4	M95580	Yeast	ORF1	C	442	29	1.9	80442	49	724D18	724D18 Arabidops
C 370	29	1.9	2483	7	YSCSEOCAT	U63062	Dictyostell		C	443	29	1.9	80970	7	AB006697	AB006697 Arabidops
C 371	29	1.9	2730	33	DDU63062	AF054907	Neutrosor		C	444	29	1.9	81542	7	AB026647	AB026647 Arabidops
C 372	29	1.9	2735	8	AF054907	U17641	Neutrosor		C	445	29	1.9	82415	7	AB005244	AB005244 Arabidops
C 373	29	1.9	2855	7	NC017641	AF006304	Saccharom		C	446	29	1.9	82571	55	AC025116	AC025116 Mus musculu
C 374	29	1.9	3069	8	AF006304	AF081801	Dictyoste		C	447	29	1.9	82603	43	AC025116	AC025116 Mus musculu
C 375	29	1.9	3386	7	ACA9670	AF109776	Metapene		C	448	29	1.9	83097	7	AB025624	AB025624 Arabidops
C 376	29	1.9	3625	34	AF109776	Z36019	S.cerevisia		C	449	29	1.9	83535	69	AC012619	AC012619 Homo sapi
C 377	29	1.9	3705	34	AF109776	J05161	Earthworm		C	450	29	1.9	83544	7	AB025615	AB025615 Arabidops
C 378	29	1.9	4009	7	SCYBR150C	M61742	A.thaliana		C	451	29	1.9	83957	10	HS885817	HS885817 Arabidops
C 379	29	1.9	4037	33	LDMHBC	Z74796	S.cerevisia		C	452	29	1.9	85992	7	AB009052	AB009052 Arabidops
C 380	29	1.9	4037	33	LDMHBC	D00319	S.cerevisia		C	453	29	1.9	86599	7	AB009050	AB009050 Arabidops
C 381	29	1.9	4418	7	ATHPATPC	U03477	Drosophila		C	454	29	1.9	86801	49	AC007190	AC007190 Genomic s
C 382	29	1.9	4434	7	SCYLOD55C	U23477	Dictyostell		C	455	29	1.9	87122	32	HSX313F4	HSX313F4 Arabidops
C 383	29	1.9	4644	81	CPVATI	Z22537	C.paryum	pr	C	456	29	1.9	87487	7	AB015474	AB015474 Arabidops
C 384	29	1.9	5081	43	AC019721	M25161	Human	NK-K-	C	457	29	1.9	87967	8	AC005223	AC005223 Arabidops
C 385	29	1.9	5577	33	DDU23477	U18844	Saccharomyc		C	458	29	1.9	88031	51	AC004378	AC004378 Drosophila
C 386	29	1.9	6030	33	CPOCYWLP	AC014339	Drosophila		C	459	29	1.9	89338	49	AC006068	AC006068 Arabidops
C 387	29	1.9	7675	8	HDMNAXTP1				C	460	29					
C 388	29	1.9	8443	8	SCE3612				C	461	29					
C 389	29	1.9	8661	42	AC014329				C	462	29					

463	29	1.9	89779	7	AB005234	AB005234 Arabidops	536	29	1.9	131547	72	AC011380	AC011380 Homo sapi
464	29	1.9	90065	34	AC005147	AC005147 Drosophill	537	29	1.9	132596	42	AC013046	AC013046 Drosophill
465	29	1.9	91480	10	HS0283F3	AL079333 Human DNA	538	29	1.9	133318	41	AC007770	AC007770 Drosophill
466	29	1.9	92153	41	AL353802	AL353802 Homo sapi	539	29	1.9	133774	53	AC016007	AC016007 Homo sapi
467	29	1.9	94972	49	AC005700	AC005700 Arabidops	540	29	1.9	144292	43	AC019051	AC019051 Homo sapi
468	29	1.9	95695	80	AC051631	AC051631 Arabidops	541	29	1.9	135038	11	H00YMXD703	L78810 Homo sapien
469	29	1.9	95933	43	AC015848	AC015848 Homo sapi	542	29	1.9	135149	55	AC018818	AC018818 Homo sapi
470	29	1.9	96075	42	AC018294	AC018294 Homo sapi	543	29	1.9	135777	56	AC024510	AC024510 Homo sapi
471	29	1.9	96902	39	AC012599	AC012599 Homo sapi	544	29	1.9	136842	76	AC027118	AC027118 Homo sapi
472	29	1.9	97263	49	ATAC013258	AC013258 Arabidops	545	29	1.9	137438	67	AC024993	AC024993 Homo sapi
473	29	1.9	98274	78	AC062019	AC062019 Homo sapi	546	29	1.9	137535	45	AC020914	AC020914 Homo sapi
474	29	1.9	99332	11	HS029818	AL096764 Human DNA	547	29	1.9	137650	41	AF161801	AF161801 Homo sapi
475	29	1.9	99357	72	AC020933	AC020933 Homo sapi	548	29	1.9	138187	60	AC011653	AC011653 Homo sapi
476	29	1.9	100000	9	AB020877	AB020877 Homo sapi	549	29	1.9	139187	12	MMNH461	MMNH461 Mus muscu
477	29	1.9	100000	9	AP000069	AP000069 Homo sapi	550	29	1.9	139480	39	H0AC002331	AC002331 Homo sapi
478	29	1.9	100000	9	AP000499	AP000499 Homo sapi	551	29	1.9	139820	39	AC005100	AC005100 Homo sapi
479	29	1.9	100364	39	AC007590	AC007590 Homo sapi	552	29	1.9	140136	40	AC017381	AC017381 Homo sapi
480	29	1.9	100835	49	ATAC016795	AC016795 Arabidops	553	29	1.9	140644	70	AC016797	AC016797 Homo sapi
481	29	1.9	100975	8	AC007980	AC007980 Arabidops	554	29	1.9	141056	40	HS1184F4	HS1184F4 Homo sapi
482	29	1.9	101371	8	AC002534	AC002534 Arabidops	555	29	1.9	141201	60	AC022637	AC022637 Homo sapi
483	29	1.9	102621	33	DMBR43E12	AL138971 Drosophill	556	29	1.9	141243	57	AC008286	AC008286 Drosophill
484	29	1.9	103225	55	AC022405	AC022405 Homo sapi	557	29	1.9	141391	72	AC024309	AC024309 Homo sapi
485	29	1.9	104607	8	ATF28J14	AL163652 Arabidops	558	29	1.9	141591	72	AC024309	AC024309 Homo sapi
486	29	1.9	105306	8	ATF9D24	AL137081 Arabidops	559	29	1.9	142651	67	HS202121	HS202121 Homo sapi
487	29	1.9	105448	50	AC010494	AC010494 Homo sapi	560	29	1.9	142807	10	HS33B19	HS33B19 Homo sapi
488	29	1.9	105733	49	AC012561	AC012561 Arabidops	561	29	1.9	143240	42	AC007812	AC007812 Homo sapi
489	29	1.9	105937	49	ATAC012679	AC012679 Arabidops	562	29	1.9	143429	11	HS309H15	HS309H15 Homo sapi
490	29	1.9	106487	12	AC007585	AC007585 Mus muscu	563	29	1.9	143450	78	AC017124	AC017124 Homo sapi
491	29	1.9	106870	42	AF212852	AF212852 Homo sapi	564	29	1.9	143459	48	AC061711	AC061711 Homo sapi
492	29	1.9	106956	42	AC009909	AC009909 Drosophill	565	29	1.9	144639	32	HS044211	HS044211 Homo sapi
493	29	1.9	107950	7	ATT17F15	AL049658 Arabidops	566	29	1.9	144784	56	AC015761	AC015761 Homo sapi
494	29	1.9	109016	8	ATT10K17	AL132977 Arabidops	567	29	1.9	145481	53	AC009338	AC009338 Drosophill
495	29	1.9	110000	31	CEV63D3_1	Continuation (2 of	568	29	1.9	145538	10	HS245G19	Z92542 Human DNA
496	29	1.9	110000	32	CEV105EB_0	AL022594 Ctenothab	569	29	1.9	145530	70	AC018572	AC018572 Homo sapi
497	29	1.9	110000	32	HSY214H10_0	AL022344 Homo sapi	570	29	1.9	145835	59	AC020560	AC020560 Homo sapi
498	29	1.9	110000	32	HSY214H10_1	Continuation (2 of	571	29	1.9	145870	56	AC023581	AC023581 Homo sapi
499	29	1.9	111524	42	AF206725	AF206725 Homo sapi	572	29	1.9	146468	62	AC010765	AC010765 Homo sapi
500	29	1.9	111768	10	HS879J18	AL035400 Homo sapi	573	29	1.9	146501	69	AC002243	AC002243 Homo sapi
501	29	1.9	111945	8	ATF1C12	AL022324 Arabidops	574	29	1.9	147765	55	AC010412	AC010412 Homo sapi
502	29	1.9	112206	51	AC008317	AC008317 Drosophill	575	29	1.9	147812	78	AC062025	AC062025 Homo sapi
503	29	1.9	112460	10	HS1104E15	AL022317 Homo sapi	576	29	1.9	147902	11	AC005670	AC005670 Homo sapi
504	29	1.9	112930	55	AC008795	AL003704 Homo sapi	577	29	1.9	148606	11	HSB451C14	AF056116 Fugu rubr
505	29	1.9	113956	40	HS74M1	AC008335 Drosophill	578	29	1.9	149337	76	AC034150	AC034150 Homo sapi
506	29	1.9	114958	51	AC008356	AC008356 Arabidops	579	29	1.9	149447	40	AL353139	AL353139 Homo sapi
507	29	1.9	116177	8	T13D8	AC005090 Homo sapi	580	29	1.9	149594	31	AP000856	AP000856 Homo sapi
508	29	1.9	116218	39	AC005090	AC005090 Homo sapi	581	29	1.9	149748	61	AC021236	AC021236 Homo sapi
509	29	1.9	116370	10	HS569M23	AL031666 Human DNA	582	29	1.9	150445	32	AL161742	AL161742 Homo sapi
510	29	1.9	117406	50	AC010792	AC010792 Homo sapi	583	29	1.9	150445	32	AL161742	AL161742 Homo sapi
511	29	1.9	118322	55	AC008891	AC008891 Homo sapi	584	29	1.9	151429	11	AC005279	AC005279 Homo sapi
512	29	1.9	118595	39	AC002379	AC002379 Human BAC	585	29	1.9	151429	11	AC005279	AC005279 Homo sapi
513	29	1.9	118893	32	AL138839	AL138839 Homo sapi	586	29	1.9	151670	56	AC022669	AC022669 Homo sapi
514	29	1.9	118931	41	AF188031	AF188031 Homo sapi	587	29	1.9	152512	72	AC025799	AC025799 Homo sapi
515	29	1.9	121290	77	AC034236	AC034236 Arabidops	588	29	1.9	152544	79	AC011309	AC011309 Homo sapi
516	29	1.9	122796	73	AC009135	AC009135 Homo sapi	589	29	1.9	152604	78	AC016671	AC016671 Homo sapi
517	29	1.9	122877	53	AC015768	AC015768 Homo sapi	590	29	1.9	153009	41	AC009801	AC009801 Homo sapi
518	29	1.9	123183	49	AC006841	AC006841 Arabidops	591	29	1.9	153098	33	PFMAL3P2	AL034538 Plasmodin
519	29	1.9	123695	11	HSJ58113	AL038826 Human DNA	592	29	1.9	153205	60	AC025828	AC025828 Homo sapi
520	29	1.9	123766	41	AC007813	AC007813 Drosophill	593	29	1.9	153234	78	AC026947	AC026947 Homo sapi
521	29	1.9	124700	34	AC007822	AC007822 Drosophill	594	29	1.9	153323	77	AC023665	AC023665 Homo sapi
522	29	1.9	124700	34	AC005558	AC005558 Drosophill	595	29	1.9	153704	69	AC020919	AC020919 Homo sapi
523	29	1.9	125644	73	AC034234	AL034234 Homo sapi	596	29	1.9	153738	74	AC040161	AC040161 Homo sapi
524	29	1.9	125900	32	AL157403	AL157403 Homo sapi	597	29	1.9	154103	76	AC051622	AC051622 Mus muscu
525	29	1.9	125957	41	AC005959	AC005959 Homo sapi	598	29	1.9	154788	10	HS325F22	AL025458 Human DNA
526	29	1.9	127335	10	HS257E24	AL034424 Human DNA	599	29	1.9	155021	39	AC007159	AC007159 Homo sapi
527	29	1.9	127867	10	HS694B14	AL031673 Human DNA	600	29	1.9	155069	78	AC018730	AC018730 Homo sapi
528	29	1.9	128117	39	AC004000	AC004000 Human PAC	601	29	1.9	155125	52	AC021069	AC021069 Homo sapi
529	29	1.9	128132	55	AC007983	AC007983 Drosophill	602	29	1.9	155323	68	AC009824	AC009824 Homo sapi
530	29	1.9	128722	10	HS15005	AL021154 Homo sapi	603	29	1.9	155450	11	AC005951	AC005951 Homo sapi
531	29	1.9	128969	10	HS204E5	Z98941 Human DNA	604	29	1.9	155526	57	AC013371	AC013371 Homo sapi
532	29	1.9	130472	32	AL188041	AL188041 Homo sapi	605	29	1.9	155537	59	AC017072	AC017072 Homo sapi
533	29	1.9	130668	80	AC005736	AC005736 Homo sapi	606	29	1.9	155795	70	AC027270	AC027270 Homo sapi
534	29	1.9	130755	11	AC004253	AL031280 Human DNA	607	29	1.9	155829	79	AC013710	AC013710 Homo sapi
535	29	1.9	131318	10	HS187B23	AL031280 Human DNA	608	29	1.9	155998	54	AC015567	AC015567 Homo sapi

609	29	1.9	156843	79	AC011145	Homo sapi	682	29	1.9	171772	42	AC009584	Homo sapi
610	29	1.9	156928	75	AC026937	Homo sapi	683	29	1.9	172104	57	AC024933	Homo sapi
611	29	1.9	157289	50	AC009234	Homo sapi	684	29	1.9	172913	53	AC002391	Homo sapi
612	29	1.9	157320	55	AC011365	Homo sapi	685	29	1.9	173116	31	AP001375	Homo sapi
613	29	1.9	157321	39	AC004912	Homo sapi	686	29	1.9	173309	61	AC025037	Homo sapi
614	29	1.9	157340	73	AC015652	Homo sapi	687	29	1.9	173781	34	AC006562	Homo sapi
615	29	1.9	157587	32	AL139125	Homo sapi	688	29	1.9	173932	71	AC013793	Homo sapi
616	29	1.9	157751	41	AC007098	Homo sapi	689	29	1.9	174090	54	AC002130	Homo sapi
617	29	1.9	157776	32	AL138893	Homo sapi	690	29	1.9	174517	76	AC009927	Homo sapi
618	29	1.9	158213	73	AC015669	Homo sapi	691	29	1.9	174731	42	AC027031	Homo sapi
619	29	1.9	158441	60	AC016803	Homo sapi	692	29	1.9	174732	42	AC012160	Homo sapi
620	29	1.9	158623	54	AC010931	Homo sapi	693	29	1.9	174977	69	AC022608	Homo sapi
621	29	1.9	158893	42	AC008031	Typhoso	694	29	1.9	175167	72	AC011404	Homo sapi
622	29	1.9	159270	41	AC007241	Homo sapi	695	29	1.9	175441	54	AC022869	Homo sapi
623	29	1.9	159934	31	AP001330	Homo sapi	696	29	1.9	175516	60	AC006380	Homo sapi
624	29	1.9	161136	32	AL138721	Homo sapi	697	29	1.9	175657	74	AC027663	Homo sapi
625	29	1.9	161739	78	AC018868	Homo sapi	698	29	1.9	176140	77	AC034172	Homo sapi
626	29	1.9	161778	56	AC024892	Homo sapi	699	29	1.9	176186	39	AC006825	Homo sapi
627	29	1.9	161930	75	AC026828	Homo sapi	700	29	1.9	176327	75	AC018859	Homo sapi
628	29	1.9	161955	77	AC034100	Mus muscu	701	29	1.9	176424	31	AP000923	Homo sapi
629	29	1.9	162013	11	CNS01DMQ		702	29	1.9	176595	40	CNS05TBN	
630	29	1.9	162481	8	AF061282	Sorghum b	703	29	1.9	176612	57	AC021308	Homo sapi
631	29	1.9	162481	61	AC018954	Homo sapi	704	29	1.9	176895	55	AC010327	Homo sapi
632	29	1.9	162532	78	AC021014	Homo sapi	705	29	1.9	177802	59	AC018718	Homo sapi
633	29	1.9	162636	43	AC009551	Homo sapi	706	29	1.9	178181	41	AL353684	Homo sapi
634	29	1.9	162822	11	AC005751	Homo sapi	707	29	1.9	178477	61	AC022801	Homo sapi
635	29	1.9	162925	70	AC018528	Homo sapi	708	29	1.9	178683	40	AL136129	Homo sapi
636	29	1.9	163062	71	AC022352	Oryza sat	709	29	1.9	179798	79	AC015978	Homo sapi
637	29	1.9	163277	76	AC010095	Homo sapi	710	29	1.9	179820	69	AC011679	Homo sapi
638	29	1.9	163373	74	AC024263	Homo sapi	711	29	1.9	179884	40	AL161896	Homo sapi
639	29	1.9	163420	56	AC012576	Homo sapi	712	29	1.9	180046	42	AL161882	Homo sapi
640	29	1.9	163779	68	AC026670	Homo sapi	713	29	1.9	180112	68	AC013464	Homo sapi
641	29	1.9	163780	75	AC025806	Homo sapi	714	29	1.9	180385	68	AC012184	Homo sapi
642	29	1.9	163880	41	AC007597	Homo sapi	715	29	1.9	180485	69	AC027145	Homo sapi
643	29	1.9	163907	41	AC011114	Homo sapi	716	29	1.9	180737	70	AC026472	Homo sapi
644	29	1.9	164138	44	AC012593	Homo sapi	717	29	1.9	180789	44	AC010683	Homo sapi
645	29	1.9	164399	33	PF041236		718	29	1.9	180919	39	AF088219	Homo sapi
646	29	1.9	164437	32	AL137848	Homo sapi	719	29	1.9	181060	61	AC021037	Homo sapi
647	29	1.9	164548	31	CE07002	Caenorhab	720	29	1.9	181150	78	AC017020	Homo sapi
648	29	1.9	164945	61	AC024968	Homo sapi	721	29	1.9	181183	42	AC013288	Homo sapi
649	29	1.9	164945	61	AC024968	Homo sapi	722	29	1.9	181313	61	AC021963	Homo sapi
650	29	1.9	165000	40	AL138895	Homo sapi	723	29	1.9	181466	79	AC005073	Homo sapi
651	29	1.9	166276	54	AC019239	Homo sapi	724	29	1.9	181541	59	AC013445	Homo sapi
652	29	1.9	166292	42	AC013271	Homo sapi	725	29	1.9	182049	68	AC007923	Homo sapi
653	29	1.9	166335	70	AC011946	Homo sapi	726	29	1.9	182798	69	AC025394	Homo sapi
654	29	1.9	166416	53	AC016866	Homo sapi	727	29	1.9	183082	74	AC010138	Homo sapi
655	29	1.9	166445	78	AC009898	Homo sapi	728	29	1.9	183210	40	AL162742	Homo sapi
656	29	1.9	166701	11	AC005747	Homo sapi	729	29	1.9	183575	40	AL161648	Homo sapi
657	29	1.9	166807	78	AC008060	Homo sapi	730	29	1.9	183837	57	AC022813	Homo sapi
658	29	1.9	166832	41	AC011830	Homo sapi	731	29	1.9	183855	56	AC022770	Homo sapi
659	29	1.9	166892	39	AC007463	Homo sapi	732	29	1.9	183963	60	AC022430	Homo sapi
660	29	1.9	166942	10	HS1054A22		733	29	1.9	184031	69	AC020782	Homo sapi
661	29	1.9	166944	40	AL353072	Homo sapi	734	29	1.9	184040	71	AC011877	Homo sapi
662	29	1.9	167082	53	AC021472	Homo sapi	735	29	1.9	184399	53	AC020573	Homo sapi
663	29	1.9	167722	44	AC021243	Homo sapi	736	29	1.9	184460	71	AC020677	Homo sapi
664	29	1.9	167825	11	CNS0180U		737	29	1.9	184505	32	AL139190	Homo sapi
665	29	1.9	168468	39	AC005510	Homo sapi	738	29	1.9	184521	54	AC016297	Homo sapi
666	29	1.9	168481	67	AC025094	Homo sapi	739	29	1.9	184949	39	AC005029	Homo sapi
667	29	1.9	168620	43	AC015711	Homo sapi	740	29	1.9	185019	54	AC002131	Homo sapi
668	29	1.9	168674	61	AC021732	Homo sapi	741	29	1.9	185249	11	CNS01DTM	
669	29	1.9	168837	79	AC007629	Homo sapi	742	29	1.9	185461	31	AP001107	Homo sapi
670	29	1.9	169287	40	AL157781	Homo sapi	743	29	1.9	185471	59	AC019208	Homo sapi
671	29	1.9	169308	78	AC041047	Homo sapi	744	29	1.9	186087	72	AC026003	Homo sapi
672	29	1.9	169443	32	AL139809	Homo sapi	745	29	1.9	186212	40	AL161616	Homo sapi
673	29	1.9	170046	40	AL162586	Homo sapi	746	29	1.9	187345	72	AC025924	Homo sapi
674	29	1.9	170561	67	AC008690	Homo sapi	747	29	1.9	187345	54	AC011076	Homo sapi
675	29	1.9	170894	79	AC007404	Homo sapi	748	29	1.9	187396	32	AL157758	Homo sapi
676	29	1.9	170896	70	AC011010	Homo sapi	749	29	1.9	187883	39	AC008268	Homo sapi
677	29	1.9	170970	63	AC021564	Homo sapi	750	29	1.9	188080	42	AC012195	Homo sapi
678	29	1.9	171005	55	AC018495	Homo sapi	751	29	1.9	188735	11	CNS00MBM	
679	29	1.9	171057	44	AC012470	Homo sapi	752	29	1.9	189625	63	AC021547	Homo sapi
680	29	1.9	171398	60	AC022060	Homo sapi	753	29	1.9	190159	40	AL137248	Homo sapi
681	29	1.9	171532	78	AC016993	Homo sapi	754	29	1.9	190650	60	AC013799	Homo sapi

755	29	1.9 191060	60	AC021468	Homo sapi	AC021468	828	29	1.9 259920	58	AE003828	AE003828 Drosophila
756	29	1.9 192366	34	AC008369	Drosophila	AC008369	829	29	1.9 260067	58	AE003807	AE003807 Drosophila
757	29	1.9 192526	68	AC021138	Homo sapi	AC021138	830	29	1.9 260104	34	AE003687	AE003687 Drosophila
758	29	1.9 192590	39	AC009509	Homo sapi	AC009509	831	29	1.9 260104	34	AE003687	AE003687 Drosophila
759	29	1.9 192791	42	AC009364	Homo sapi	AC009364	832	29	1.9 264996	40	AL137016	AL137016 Homo sapi
760	29	1.9 192828	78	AC017036	Homo sapi	AC017036	833	29	1.9 264996	40	AL137016	AL137016 Homo sapi
761	29	1.9 193224	78	AC025668	Mus muscu	AC025668	834	29	1.9 265562	70	AC027308	AC027308 Homo sapi
762	29	1.9 193224	78	AC025668	Mus muscu	AC025668	835	29	1.9 267770	53	AC016745	AC016745 Homo sapi
763	29	1.9 193881	32	AL157386	Homo sapi	AL157386	836	29	1.9 268100	42	AC018643	AC018643 Homo sapi
764	29	1.9 193930	79	AC017111	Homo sapi	AC017111	837	29	1.9 269082	31	CE7642	CE7642 Homo sapi
765	29	1.9 193981	76	AC016826	Homo sapi	AC016826	838	29	1.9 269542	53	AC021155	AC021155 Homo sapi
766	29	1.9 194137	31	AP001446	Homo sapi	AP001446	839	29	1.9 269661	54	AC008482	AC008482 Homo sapi
767	29	1.9 194317	40	AL139824	Homo sapi	AL139824	840	29	1.9 270000	10	AB026898	AB026898 Homo sapi
768	29	1.9 194562	68	AC025408	Homo sapi	AC025408	841	29	1.9 272476	60	AC025809	AC025809 Homo sapi
769	29	1.9 194573	44	AC016914	Homo sapi	AC016914	842	29	1.9 275185	54	AC008478	AC008478 Homo sapi
770	29	1.9 194841	71	AC021139	Homo sapi	AC021139	843	29	1.9 281502	34	AE003546	AE003546 Drosophila
771	29	1.9 194905	68	AC009517	Homo sapi	AC009517	844	29	1.9 289755	51	AC020664	AC020664 Homo sapi
772	29	1.9 195470	43	AC022259	Homo sapi	AC022259	845	29	1.9 295312	34	AE003582	AE003582 Drosophila
773	29	1.9 195653	39	AC005873	clb_15_o	AC005873	846	29	1.9 298536	34	AE003504	AE003504 Drosophila
774	29	1.9 196993	32	AL158829	Homo sapi	AL158829	847	29	1.9 300205	34	AE003570	AE003570 Drosophila
775	29	1.9 197045	76	AC016821	Homo sapi	AC016821	848	29	1.9 303648	34	AE003425	AE003425 Drosophila
776	29	1.9 197071	52	AC020788	Homo sapi	AC020788	849	29	1.9 303648	34	AE003486	AE003486 Drosophila
777	29	1.9 197357	70	AC019049	Homo sapi	AC019049	850	29	1.9 311662	32	AL136181	AL136181 Homo sapi
778	29	1.9 197540	72	AC023555	Homo sapi	AC023555	851	29	1.9 325686	43	AC021094	AC021094 Homo sapi
779	29	1.9 197791	32	AL136178	Homo sapi	AL136178	852	29	1.9 329391	34	AE003681	AE003681 Homo sapi
780	29	1.9 198427	8	ATCHRIV52	Arabidops	AL161552	853	29	1.9 335282	53	AC017085	AC017085 Homo sapi
781	29	1.9 198948	32	AL136380	Homo sapi	AL136380	854	28	1.8 118	14	SCPI04123	SCPI04123 Homo sapi
782	29	1.9 199065	32	AL136380	Homo sapi	AL136380	855	28	1.8 161	8	AF027616	AF027616 Arabidops
783	29	1.9 199612	55	AC016543	Homo sapi	AC016543	856	28	1.8 260	12	RN089598	RN089598 Rattus norv
784	29	1.9 199957	40	AC012043	Homo sapi	AC012043	857	28	1.8 277	13	G09999	G09999 human STS C
785	29	1.9 200746	44	AL133343	Homo sapi	AL133343	858	28	1.8 301	13	G41853	G41853 human STS C
786	29	1.9 201990	39	AC002385	Human BAC	AC002385	859	28	1.8 344	13	G10326	G10326 human STS C
787	29	1.9 202192	55	AC023762	Homo sapi	AC023762	860	28	1.8 355	13	HM071995	HM071995 human STS C
788	29	1.9 202223	11	AC003664	Homo sapi	AC003664	861	28	1.8 410	13	G47862	G47862 human STS C
789	29	1.9 202551	31	AP001484	Homo sapi	AP001484	862	28	1.8 433	13	HM078001	HM078001 human STS U
790	29	1.9 203226	43	AC017002	Homo sapi	AC017002	863	28	1.8 527	34	AF140072	AF140072 Apis mell
791	29	1.9 203228	40	AL135903	Homo sapi	AL135903	864	28	1.8 574	13	G40358	G40358 human STS U
792	29	1.9 203418	75	AC026011	Homo sapi	AC026011	865	28	1.8 706	34	AF108931	AF108931 Mercenari
793	29	1.9 204617	70	AC025644	Homo sapi	AC025644	866	28	1.8 858	13	CNS011VE	CNS011VE Homo sapi
794	29	1.9 204885	32	AL137249	Homo sapi	AL137249	867	28	1.8 900	103	CNS011ETN	CNS011ETN Homo sapi
795	29	1.9 205685	44	AC013446	Homo sapi	AC013446	868	28	1.8 979	13	CNS011G25	CNS011G25 Homo sapi
796	29	1.9 206476	52	AC012557	Homo sapi	AC012557	869	28	1.8 986	7	SASCHS1	SASCHS1 Homo sapi
797	29	1.9 207375	32	CNS0048W	Homo sapi	AL080279	870	28	1.8 1004	13	CNS011FE6	CNS011FE6 Homo sapi
798	29	1.9 207608	76	AC022397	Homo sapi	AC022397	871	28	1.8 1036	13	CNS01106	CNS01106 Homo sapi
799	29	1.9 207722	79	AC0192306	Homo sapi	AC0192306	872	28	1.8 1310	15	AX002276	AX002276 Sequence
800	29	1.9 207869	69	AC023050	Homo sapi	AC023050	873	28	1.8 1675	33	BMRDNTAT	BMRDNTAT Homo sapi
801	29	1.9 208363	11	AC003101	Homo sapi	AC003101	874	28	1.8 2005	33	SCSMOXIT	SCSMOXIT Homo sapi
802	29	1.9 208673	68	AC026100	Homo sapi	AC026100	875	28	1.8 2495	34	AF212312	AF212312 Homo sapi
803	29	1.9 208949	53	AC024171	Homo sapi	AC024171	876	28	1.8 2873	7	SASCHS2	SASCHS2 Homo sapi
804	29	1.9 209382	39	AC005071	Homo sapi	AC005071	877	28	1.8 3000	49	AF163819	AF163819 Arabidops
805	29	1.9 209787	44	AC012493	Homo sapi	AC012493	878	28	1.8 3263	15	AX002278	AX002278 Sequence
806	29	1.9 209872	57	AC012301	Homo sapi	AC012301	879	28	1.8 3324	10	HSR4270993	HSR4270993 Homo sapi
807	29	1.9 210829	44	AC012506	Homo sapi	AC012506	880	28	1.8 3904	7	SCYDR080W	SCYDR080W Homo sapi
808	29	1.9 211188	55	AC024093	Homo sapi	AC024093	881	28	1.8 4001	34	AF0670153	AF0670153 Homo sapi
809	29	1.9 211892	55	AC008756	Homo sapi	AC008756	882	28	1.8 4273	2	AF012911	AF012911 Plasmid p
810	29	1.9 213813	73	AC007610	Homo sapi	AC007610	883	28	1.8 4286	15	AX002277	AX002277 Sequence
811	29	1.9 215574	40	AL157710	Homo sapi	AL157710	884	28	1.8 4671	5	A58998	A58998 Sequence 2
812	29	1.9 218094	43	AC015797	Mus muscu	AC015797	885	28	1.8 4742	81	AF015297	AF015297 Human her
813	29	1.9 218741	73	AC008869	Homo sapi	AC008869	886	28	1.8 4985	14	AF147464	AF147464 T7 expires
814	29	1.9 219063	53	AC016943	Homo sapi	AC016943	887	28	1.8 5789	7	AB000223	AB000223 Yeast DNA
815	29	1.9 219151	73	AC012175	Homo sapi	AC012175	888	28	1.8 7249	42	AC017839	AC017839 Drosophila
816	29	1.9 223906	40	AL162501	Homo sapi	AL162501	889	28	1.8 8733	34	AF212313	AF212313 Drosophila
817	29	1.9 226905	58	AE003773	Drosophila	AE003773	890	28	1.8 11572	34	AF177386	AF177386 Drosophila
818	29	1.9 230081	52	AC009387	Homo sapi	AC009387	891	28	1.8 12351	42	AF146362	AF146362 Drosophila
819	29	1.9 240257	41	HSBA416N2	Homo sapi	AL119299	892	28	1.8 12575	54	LPIC04502	LPIC04502 Drosophila
820	29	1.9 244951	32	AL139084	Homo sapi	AL139084	893	28	1.8 13965	81	AF015298	AF015298 Human her
821	29	1.9 246856	53	AC013242	Homo sapi	AC013242	894	28	1.8 19451	42	AC014476	AC014476 Drosophila
822	29	1.9 248778	41	MAE009774	Homo sapi	MAE009774	895	28	1.8 21951	81	HR013194	HR013194 Human herpe
823	29	1.9 250611	12	MAE0090663	Homo sapi	MAE0090663	896	28	1.8 27621	73	AC020338	AC020338 Drosophila
824	29	1.9 251384	41	HS1191N16	Homo sapi	HS1191N16	897	28	1.8 32821	47	SCCHRO1V	SCCHRO1V Homo sapi
825	29	1.9 251726	69	AC008737	Homo sapi	AC008737	898	28	1.8 33455	33	CEW0448	CEW0448 Caenorhabdi
826	29	1.9 252455	41	AC009260	Homo sapi	AC009260	899	28				
827	29	1.9 253747	34	AE003723	Drosophila	AE003723	900	28				

901	28	1.8	34270	9	AP000270	Homo sapi	974	28	1.8	164563	67	AC025093	Homo sapi
902	28	1.8	34721	34	AC004309	Drosophila	975	28	1.8	167957	68	AC011205	Homo sapi
903	28	1.8	38760	7	SC8554	Z46796 S. cerevisia	976	28	1.8	169131	57	AC013780	Homo sapi
904	28	1.8	39612	42	AC015516	Homo sapi	977	28	1.8	169623	32	AL158162	Homo sapi
905	28	1.8	43125	34	AF147779	Drosophila	978	28	1.8	170703	67	AC008476	Homo sapi
906	28	1.8	43843	11	AC005796	Homo sapi	979	28	1.8	170948	72	AC012312	Homo sapi
907	28	1.8	45284	77	AC055745	Homo sapi	980	28	1.8	172830	77	AC027032	Homo sapi
908	28	1.8	62091	31	DMER8018	Homo sapi	981	28	1.8	173570	56	AC023261	Homo sapi
909	28	1.8	64160	51	AC022984	Homo sapi	982	28	1.8	176843	76	AC026555	Homo sapi
910	28	1.8	66146	43	AC020262	Drosophila	983	28	1.8	177143	69	AC022440	Homo sapi
911	28	1.8	66730	49	AC005561	Arabidops	984	28	1.8	177319	43	AC021651	Homo sapi
912	28	1.8	67351	70	AC027584	Homo sapi	985	28	1.8	178204	74	AC011028	Homo sapi
913	28	1.8	71320	43	AC020124	Drosophila	986	28	1.8	179026	60	AC018857	Homo sapi
914	28	1.8	72887	42	AC017882	Homo sapi	987	28	1.8	179637	41	AC015571	Homo sapi
915	28	1.8	73367	43	AC021929	Homo sapi	988	28	1.8	180043	54	AC023081	Homo sapi
916	28	1.8	73677	43	AC021929	Homo sapi	989	28	1.8	180315	39	AC007773	Homo sapi
917	28	1.8	76018	56	AC008698	Homo sapi	990	28	1.8	181629	31	AP000853	Homo sapi
918	28	1.8	77483	7	AB028607	Arabidops	991	28	1.8	181286	78	AC026988	Homo sapi
919	28	1.8	81637	42	AC015355	Drosophila	992	28	1.8	185180	40	AL136131	Homo sapi
920	28	1.8	83689	7	AB023608	Arabidops	993	28	1.8	185706	41	AL136131	Homo sapi
921	28	1.8	87104	8	AT15C23	AB049500 Arabidops	994	28	1.8	186256	32	AL158033	Homo sapi
922	28	1.8	87835	7	AB005237	Arabidops	995	28	1.8	186463	67	AC024196	Homo sapi
923	28	1.8	92620	7	AB026636	Arabidops	996	28	1.8	186810	55	AC008734	Homo sapi
924	28	1.8	93639	49	U90439	U90439 Arabidops	997	28	1.8	186937	73	AP001319	Homo sapi
925	28	1.8	98253	8	AC002423	Genomic s	998	28	1.8	188599	31	AP001319	Homo sapi
926	28	1.8	100000	9	AP000032	Homo sapi	999	28	1.8	189219	54	AC009680	Homo sapi
927	28	1.8	100000	9	AP000104	Homo sapi	1000	28	1.8	189219	54	AC009680	Homo sapi
928	28	1.8	100000	9	AP000180	Homo sapi							
929	28	1.8	100310	31	F17A13	AL096692 Arabidops							
930	28	1.8	101158	8	AC002560	Arabidops							
931	28	1.8	102460	11	AC005857	Homo sapi							
932	28	1.8	105907	67	AC026429	Homo sapi							
933	28	1.8	108554	39	AC005226	Homo sapi							
934	28	1.8	109692	73	AC008683	Homo sapi							
935	28	1.8	110000	31	CEY37H9.1	Continuation (2 of							
936	28	1.8	110000	32	CEY71A12.2	Continuation (3 of							
937	28	1.8	111109	10	HS475N16	AL033587 Human DNA							
938	28	1.8	111109	10	HS475N16	AL033587 Human DNA							
939	28	1.8	115837	53	AC009736	AL020298 Drosophila							
940	28	1.8	115863	10	HS268D13	AL020298 Drosophila							
941	28	1.8	116370	31	AP000683	AP000683 Homo sapi							
942	28	1.8	117752	39	AC006205	AC006205 Homo sapi							
943	28	1.8	117859	41	HS5355E2	AL108811 Homo sapi							
944	28	1.8	120159	55	AC010601	AC010601 Homo sapi							
945	28	1.8	121256	41	AC000789	AP000789 Drosophila							
946	28	1.8	122667	31	AP000839	AP000839 Homo sapi							
947	28	1.8	122709	73	AC008722	AC008722 Homo sapi							
948	28	1.8	128265	43	AC007826	AC007826 Drosophila							
949	28	1.8	131474	55	AC008660	AC008660 Homo sapi							
950	28	1.8	131594	32	DMER17J10	AL122026 Drosophila							
951	28	1.8	134226	81	ITHICG	AL122026 Drosophila							
952	28	1.8	137040	43	AC007827	AC007827 Drosophila							
953	28	1.8	138705	31	AP000485	AP000485 Homo sapi							
954	28	1.8	139356	69	AC008750	AC008750 Homo sapi							
955	28	1.8	141475	11	AC005287	AC005287 Arabidops							
956	28	1.8	143186	49	HS29C18	297192 Human DNA							
957	28	1.8	143701	10	HS29C18	297192 Human DNA							
958	28	1.8	148606	43	AC020298	AC020298 Drosophila							
959	28	1.8	148606	43	AC020298	AC020298 Drosophila							
960	28	1.8	150269	80	AC034186	AL121875 Human DNA							
961	28	1.8	151024	57	AC016855	AC016855 Homo sapi							
962	28	1.8	151243	31	AP001208	AP001208 Homo sapi							
963	28	1.8	153773	39	HDAC004020	HDAC004020 Homo sapi							
964	28	1.8	154732	53	AC023482	AC023482 Homo sapi							
965	28	1.8	154732	53	AC023482	AC023482 Homo sapi							
966	28	1.8	155318	10	HS791K14	AL035685 Human DNA							
967	28	1.8	157851	72	AC034471	AC034471 Homo sapi							
968	28	1.8	158171	71	AC022258	AC022258 Homo sapi							
969	28	1.8	158866	72	AC025970	AC025970 Homo sapi							
970	28	1.8	159321	81	HHV6AGNM	X83413 Human herpe							
971	28	1.8	163062	71	AC022352	AC022352 Oryza sat							
972	28	1.8	163254	40	AL158212	AL158212 Homo sapi							
973	28	1.8	164079	67	AC026319	AC026319 Homo sapi							

ALIGNMENTS

RESULT	1	VERICR	2100 bp	DNA	BCT	26-Apr-1993
LOCUS	VERICR	Yersinia pestis lcrG, lcrV, and lcrH genes, complete cds.				
DEFINITION	Yersinia pestis lcrG, lcrV, and lcrH genes, complete cds.					
ACCESSION	M26405.1	GI:155448				
VERSION	M26405.1	GI:155448				
KEYWORDS	lcrG protein; lcrV protein; lcrV protein; V antigen.					
SOURCE	75kb Yersinia plasmid.					
ORGANISM	Yersinia pestis					
REFERENCE	1 (bases 1 to 2100)					
AUTHORS	Price, S.B., Leung, K.Y., Barve, S.S. and Straley, S.C.					
TITLE	Molecular analysis of lcrGVH, the V antigen operon of Yersinia					
JOURNAL	J. Bacteriol. 171, 5646-5653 (1989)					
MEDLINE	90008606					
COMMENT	Computer readable copy of sequence [J. Bacteriol. (1989) in press] kindly submitted by Price, S.					
FEATURES	07-Aug-1989.					
source	Location/Qualifiers					
	1..2100					
	/organism="Yersinia pestis"					
	/strain="K1ms"					
	/db_xref="taxon:632"					
	147..152					
	-35_signal					
	174..179					
	192..479					
	gene					
	/gene="lcrG"					
	192..479					
	gene					
	/gene="lcrG"					
	192..479					
	gene					
	/gene="lcrV"					
	463..466					
	gene					
	/gene="lcrG"					
	463..466					
	gene					
	/gene="lcrV"					

```

CDS
481..1461
/gene="lcrv"
/codon_start=1
/transl_table=11
/protein_id="AAA27641.1"
/db_xref="GI:155450"
/translation="MIRAEONPHFIEDLEKRVQULTGSSVLEELVOLVDKDKI
DISIKYDRKSEVFNANVITDIDELKRIAYFLPEDAILKGHYDQLONGIRK
EFLSSPTOMELRAFMVMEHSLADRIDDLKVIYDSMHGHDARSKEILAEI
TAEIKIYVIOEIKNHLSSSGTINHKRINLMKNIYGTDEEIFRASAERYILEK
MPTTIOYDSEKRIKIVISKDFGSENKRGATGNKNSYKKNENSHRPTTSDSK
SRPLNDIVSOKTQLOLSDITSRNSAIEALNRFQIXDYDMLDITSGK"
RBS
1463..1468
/gene="lcrh"
/gene="lcrh"
1465..1980
/gene="lcrh"
1474..1980
/gene="lcrh"
/codon_start=1
/transl_table=11
/protein_id="AAA27642.1"
/db_xref="GI:155451"
/translation="MQQETTDQEQVQLAMESFLKGGCTIAMLNEISSDTLELISLAF
NTOYSGKEDHRYQALCVLDHYDSRFFLGGACROAMGODLAISHSYGALMDIK
EPREPHAECLLORKEGLAEESGLFLAOELIADTKERKELSTRVSMLEAIKDKEM
EHECVNDP"
BASE COUNT      688 a      423 c      461 g      528 t
ORIGIN
Query Match      59.4%; Score 930; DB 1; Length 2100;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1186 TATACAGATGAGACATTTTAAAGCCAGCAGAGTACAAAATTCGAGAAAATGCT 1245
DB 1081 TATACAGATGAGACATTTTAAAGCCAGCAGAGTACAAAATTCGAGAAAATGCT 1140
OY 1246 CAACCCACATTCAGGTGATGGAGCGAGAAAATAATAGTCTCGATAAGACATTCCT 1305
DB 1141 CAACCCACATTCAGGTGATGGAGCGAGAAAATAATAGTCTCGATAAGACATTCCT 1200
OY 1306 GGAAGGAGAAATTAAGAAACCCGGCGCTGGGTAATCTGAAAAAGTCATCTTATAT 1365
DB 1201 GGAAGGAGAAATTAAGAAACCCGGCGCTGGGTAATCTGAAAAAGTCATCTTATAT 1260
OY 1366 AAAGATTAATTAATTAATTAATTCATCTTGGCACACCTGCTCGATTAAGCCAGCCGCTC 1425
DB 1261 AAAGATTAATTAATTAATTAATTCATCTTGGCACACCTGCTCGATTAAGCCAGCCGCTC 1320
OY 1426 AACGACTTGTTAGCCAAAACCACTGCTGATTAATTAATTAATTAATTAATTAATTA 1485
DB 1321 AACGACTTGTTAGCCAAAACCACTGCTGATTAATTAATTAATTAATTAATTAATTA 1380
OY 1486 GCTATTGAAGCACTGAAACCGCTTCAATTCAGAAATATGATTCAGTATGACAGCTGCTGA 1545
DB 1381 GCTATTGAAGCACTGAAACCGCTTCAATTCAGAAATATGATTCAGTATGACAGCTGCTGA 1440
OY 1546 GATGACAGCTGCTGTAATGA 1566
DB 1441 GATGACAGCTGCTGTAATGA 1461

RESULT 2
LOCUS YPCD1 70305 bp DNA BCT 22-MAR-2000
DEFINITION Yersinia pestis plasmid pCD1.
ACCESSION AL117189
VERSION AL117189.1 GI:5832423
KEYWORDS chaperone; cytotoxic effector; IS100; IS1616; IS1617; lcr;
low-calcium response; ysc; targeted effector; transposase; type III
secretion; V antigen; virulence; ylp; yop; ysc.
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis; gamma subdivision; Enterobacteriaceae;
Bacteria; Proteobacteria; Yersinia.
REFERENCE 1 (bases 1 to 70305)
AUTHORS Karlyshev, A.V. and Wren, B.W.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 70305)
AUTHORS Baker, S.G. and Mungall, K.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 70305)
AUTHORS James, R.D., Parkhill, J., Barrell, B.G. and Randal, M.A.
JOURNAL Direct Submission
TITLE Submitted (03-SEP-1999) Yersinia pestis sequencing project, Sanger
Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA
E-mail: barrel@sanger.ac.uk DNA supplied by Dr. Andrey Karlyshev
and Prof. Brendan Wren, [3]. Department of Infectious and Tropical
Diseases, London School of Hygiene and Tropical Medicine, Keppel
Street, London WC1E 7HT
NOTES:
Yersinia pestis sequencing at The Sanger Centre is funded by
Beowulf Genomics.
Details of Y. pestis sequencing at the Sanger Centre are available
on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/Y_pestis/)
CDS are numbered using the following system eg YPCP1.01c. YP (Y.
pestis), PCP1 (plasmid name), .01 (first CDS), c (complementary
strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given

```

where these have been used to deduce the initiation codon.
 CAUTION: We may not have predicted the correct initiation codon.
 Where possible we choose an initiation codon (atg, gtg, ttg or
 (att)) which is preceded by an upstream ribosome binding site
 sequence (optimally 5-13bp before the initiation codon). If this
 cannot be identified we choose the most upstream initiation codon.

Location/Qualifiers

FEATURES

source

1. 70305
 /organism="Yersinia pestis"
 /plasmid="pCD1"
 /strain="CO-92 Biovar Orientalis"
 /db_xref="taxon:632"

repeat_unit

1. 1956
 /note="IS100 element"

gene

88. 1110
 /gene="YPCD1.01"
 88. 1110
 /gene="YPCD1.01"

CDS

/gene="YPCD1.01", probable transposase, len: 340 aa;
 /note="YPCD1.01, probable transposase, len: 340 aa;
 putative insertion sequence IS100, identical to
 corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946,
 AF074612) (340 aa), fasta scores: opt: 2328 z-score:
 3808.9 E(): 0, 100.0% identity in 340 aa overlap. Similar
 to many others e.g. TRA0_ECOLI (EMBL:X14793), ISca, E.coli
 transposase for insertion sequence element IS21 (390 aa)
 (33.1% identity in 329 aa overlap). Contains Pfam match to
 entry PF00239 recombinase, site-specific recombinases,
 score 25.70, E-value 4.8e-06. Contains probable
 helix-turn-helix motif at aa 19-40 (Score 2045, +6.15 SD)"
 /codon_start=1
 /transl_table=11
 /label=YPCD1.01
 /product="putative transposase"
 /protein_id="CAB54878.1"
 /db_xref="GI:5832424"

misc_feature

112. 195
 /gene="YPCD1.01"
 /note="Pfam match to entry PF00239 recombinase,
 site-specific recombinases, score 25.70, E-value 4.8e-06"

gene

1110. 1889
 /gene="YPCD1.02"
 1110. 1889
 /gene="YPCD1.02"

CDS

/gene="YPCD1.02", probable transposase, len: 259 aa;
 /note="YPCD1.02, probable transposase, len: 259 aa;
 putative insertion sequence IS100, identical to
 corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946,
 AF074612) (259 aa), fasta scores: opt: 1658 z-score:
 2693.4 E(): 0, 100.0% identity in 259 aa overlap. Similar
 to many others e.g. ISTB_ECOLI (EMBL:X14793), ISB, E.coli
 transposase for insertion sequence element IS21 (265 aa)
 (47.4% identity in 249 aa overlap). Contains P500017
 ATP/GTP-binding site motif A (P-loop)"
 /codon_start=1
 /transl_table=11
 /label=YPCD1.02
 /product="putative ATP-binding protein"
 /protein_id="CAB54879.1"
 /db_xref="GI:5832425"

misc_feature

1434. 1457
 /gene="YPCD1.02"
 /note="P500017 ATP/GTP-binding site motif A (P-loop)"

misc_feature

1956. 46271
 /note="identical to Y.pestis KIM5 plasmid pCD1
 (EMBL:AF053946) from 46489 to 1955, and to Y.pestis KIM5
 plasmid pCD1 (EMBL:AF074612) from 59097 to 14563, except
 where noted"

gene

complement(1956..2204)
 /gene="YPCD1.03c"
 complement(1956..2204)
 /partial

CDS

/gene="YPCD1.03c", probable transposase remnant, len: 83
 aa, similar to several e.g. N-terminus of TRA0_ECOLI
 (EMBL:X14793), ISca, E.coli transposase for insertion
 sequence element IS21 (390 aa), similarity is interrupted
 by the adjacent IS100 element. The remainder of this CDS
 is in YPCD1.97c. This region is also similar to TR:068707
 (EMBL:AF053946, AF074612) Y.pestis KIM5 putative
 transposase in pCD1 (390 aa). Contains probable
 helix-turn-helix motif at aa 19-40 (Score 1833, +5.43 SD)"
 /codon_start=1
 /transl_table=11
 /label=YPCD1.03c
 /protein_id="CAB54880.1"
 /db_xref="GI:5832426"

gene

2304. 2948
 /gene="YPCD1.04"
 2304. 2948
 /gene="YPCD1.04"

CDS

/partial
 /gene="YPCD1.04", possible transposase remnant, len: 215
 aa; similar to many e.g. TR:046612 (EMBL:X78052)
 Enterobacter agglomerans IS 1222 ORFb (276 aa). Truncated
 at N-terminus"

misc_feature

2304. 2982
 /note="similar to E.agglomerans IS 1222 (EMBL:X78052) at
 DNA level"

gene

complement(3014..3406)
 /gene="YPCD1.05c"
 complement(3014..3406)
 /gene="YPCD1.05c"

CDS

/note="YPCD1.05c, syce, yera, yope targeting protein, len:
 130 aa; identical to corresponding CDS from Y.pestis KIM5
 pCD1 (EMBL:AF053946, AF074612) (130 aa), fasta scores:
 opt: 861 z-score: 1608.5 E(): 0, 100.0% identity in 130 aa
 overlap and to YERA_YERPE (EMBL:M34279) from Y.pestis
 plasmid pY019. Highly similar to TR:056910 (EMBL:Z18539),
 syce, from Yersinia enterocolitica (130 aa) (99.2%
 identity in 130 aa overlap) and YERA_YEREN (EMBL:M34278),
 yere, also from Y.enterocolitica (130 aa) (98.5% identity
 in 130 aa overlap). Similar to TR:051448 (EMBL:L27629),
 ORF1, pseudomonas aeruginosa exoenzyme S ORF1 (116 aa)
 (44.3% identity in 115 aa overlap)"
 /codon_start=1
 /transl_table=11
 /label=syce
 /product="putative yope chaperone"
 /protein_id="CAB54882.1"
 /db_xref="GI:5832428"

gene

/translation="MRSFOATITQLRQSLISIPDIIEPIVIGKVGECFCHTEHPVG
 QILMFLPSIDNNDEKETLSNHSIEQDILKPLISDVEGVGAPVLMNRPPLNSIDNS
 LITQLEMLVGGARLQTSLSIPRSFS"
 3600. 4259

Query Match	59.4%;	Score 930;	DB 2;	Length 70305;
Best Local Similarity	99.9%;	Pred. NO. 0;		
Matches 980; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1426	AAGAGCTTGGTAGGCCAAAAAACCACTAGCGTGTGATATTAATCAGCTTTAATTC	1485
Db	22075	AAGAGCTTGGTAGGCCAAAAAACCACTAGCGTGTGATATTAATCAGCTTTAATTC	22011
QY	1486	GCATTTGAAGCACTGAGCCGTTTCATTCAGAAATATGATTCAGTGCAGTCAACGCTGCTA	1545
Db	22015	GCATTTGAAGCACTGAGCCGTTTCATTCAGAAATATGATTCAGTGCAGTCAACGCTGCTA	21950
QY	1546	GATGACACGCTGTGTTAATGA	1566
Db	21935	GATGACACGCTGTGTTAATGA	21935

[illegible]

```

FEATURES
source
location/Qualifiers
1. 70504
/organism="Yersinia pestis"
/plasmid="pCD1"
/strain="KIM"
/db_xref="taxon:632"
87. 1109
/codon_start=1
/transl_table=1
/product="transposase"
/protein_id="AAC62546.1"
/db_xref="GI:2996223"
/transl_table="MVFETVMEKILHKOGMSSRAIARELGISRNTYKRYLOAKSE
PKTPRAVAVSLIDEVDYDIKRIADAHAPKIPATVATREKOGYGGMMILAFIT
SLSPDQCEPAVRETEIRPGKMOYDQKMTNRGSPPLVFAVALGTSRLATIEFDNNEH
YDLTECHRNAAFREFGVAPREVLIDNKKTVLQDAVQTGQHRPISLMQCKEMGEF
PRICRPAQTKGVERMVOYTRNSFYIPLMTRLRPMGIIVDETANHGSRMWDHVA
NORHETIQARPCDRLMEQSMALPPEKEYDVHLDENLVNDKHPHLHPPLSTYDS
FCGVA"
1109. 1888
/codon_start=1
/transl_table=1
/product="transposase"
/protein_id="AAC62557.1"
/db_xref="GI:2996234"
/transl_table="MMLQORLMALAGOLSELISNAAPLSOANDQESVYMDLE
HLHEEKLAHQKQAVYTRNAAPPAVKITEEIDFTATGAPQQLDSLRSLSTERE
ENVYLLGPSCGVGTHLAIAMGEAVRAGINVRFTADLLQSLTAORGRYKTYLOR
GVALPRLIIDEIGYTLFPOSEAKLFEQVIAKREKAMILTSMLPGWMDQTFADGAP
ALISAMIDRLIHSHVYQIGESYRLNQKKRAGVIAEANPE"
complement(1939. 2343)
/codon_start=1
/transl_table=1
/product="transposase"

```



```

gene      /protein_id="AAC62568.1"
CDS       /db_xref="GI:2996245"
           /translation="MTEYQASERRCGRINGISRSILHYCPNTPARDIPVVEVLQKLAHQ
           VYAGFGFLMFKLRQSGLPWNRVRYRRLKLNFRRRGRKRLPNRHPQPLAIDLM
           NCMWSVDMSDLTQGRPRPLNFNVVEIFVNDG"
           complement(2379. .2645)
gene      /gene="lcrs"
CDS       /complement(2379. .2645)
           /gene="lcrs"
           /codon_start=1
           /product="low calcium response locus protein S homolog"
           /protein_id="AAC62579.1"
           /db_xref="GI:2996256"
           /translation="MKKARETQQLIRVLKEVEGRHVKVDCRENGSGEASYYMKSK
           YGGMSSSDIKRNRKEEENRRLKQWYASISLDELHKVAAKL"
           complement(3193. .3540)
gene      /gene="yscm"
CDS       /complement(3193. .3540)
           /gene="yscm"
           /codon_start=1
           /translation="top proteins translocation protein M homolog"
           /protein_id="AAC62586.1"
           /db_xref="GI:2996263"
           /translation="MKINTQSLNQQTQVHGQAGRLTETNPITENSQISTAEK
           LEANRQDTAAG"
           complement(3765. .4430)
gene      /gene="yscl"
CDS       /complement(3765. .4430)
           /gene="yscl"
           /codon_start=1
           /translation="top proteins translocation protein L homolog"
           /protein_id="AAC62596.1"
           /db_xref="GI:2996273"
           /translation="MSOTQIGYAYMOPVOIIPNSLSIAGRLIRAEQYSSLTTE
           ELISAKQADLADAOEYEQKQLOQAMQADARTLOATLIHETOLQCCPFYRH
           EQQSEVVALAEKILINDYQVAMTLOVRELAIVKQYVAVNPQOALIEQI
           AKVHKDFEISYLEVTADARLDQGCILETEGIIASIDQIEALSRAISITLQMK
           VTE"
           complement(4376. .5005)
gene      /gene="ysck"
CDS       /complement(4376. .5005)
           /gene="ysck"
           /codon_start=1
           /translation="top proteins translocation protein K homolog"
           /protein_id="AAC62605.1"
           /db_xref="GI:2996282"
           /translation="MMENYITSFQLRFCPPAYVLIHQPLSLRSILPILPQMRDSAH
           LNALDEFSLDYEHPGLGALPOPOSOLELRCGLVHGSAIRVCYASLQO
           LITVNOETLQIIVQHELLIGMPPTHQMRPLPIERTIMQSLAFVLAEMEQPO
           AMKRSLRPLATPSPPVVAESQRPALQITLCHKLVQVPTCSHLK"
           complement(3005. .5739)
gene      /gene="yscj"
CDS       /complement(3005. .5739)
           /gene="yscj"
           /codon_start=1
           /translation="top proteins translocation protein J"
           /protein_id="AAC62607.1"
           /db_xref="GI:2996284"
           /translation="MKVKTSLTILITLFGCKVDIYTGISQKEGEMALIRBEG
           SADKPPDRGKIKLVESDVAAIDILIKRKYPRESTTDLVPPKGLISPIEEL
           ARLNTAKAQBLSRTSEIDGVAVARHVLPEDQNNKKGKGAASAVFIRKAAADIOF
           DTVIPIQIOLVNSIEGLAIVRIIVLPSVDVROSSHLPRNTSILTSIOVESSESGHL
           IGLSLILLPLVINIAQYFWLQKK"
           complement(5746. .6093)
gene      /gene="ysci"
CDS       /complement(5746. .6093)
           /gene="ysci"

```

```

gene      /codon_start=1
CDS       /translation="top proteins translocation protein I homolog"
           /protein_id="AAC62608.1"
           /db_xref="GI:2996285"
           /translation="MPTIEIAQADEVITITLEEAPAPPTDQIRPDANSQDGL
           GHSILKEVSDIQSKFTVKSDHRTKLAIVSNPNDMLMOWSLIRITIOEBLIAKTAG
           RMSQVETLSKRG"
           complement(6094. .6591)
gene      /gene="ysch"
CDS       /complement(6094. .6591)
           /gene="ysch"
           /codon_start=1
           /translation="top proteins translocation protein H homolog"
           /protein_id="AAC62547.1"
           /db_xref="GI:2996324"
           /translation="MTVTLNGSITSLMSSQAVSTLOPVASELKTQLENTKESAE
           KTRVLMQOYVASNPPDAVLEVATPVREALARFGQHGQSVPAIDLPRLRSVLOQ
           FDSFKRWETAILQVLEGIKPNESQVGLPYLSEILNKELMILLPSNSIYDSLNSHQ
           IDMDT"
           complement(6588. .6935)
gene      /gene="yscg"
CDS       /complement(6588. .6935)
           /gene="yscg"
           /codon_start=1
           /translation="top proteins translocation protein G homolog"
           /protein_id="AAC62548.1"
           /db_xref="GI:2996225"
           /translation="MKRYLVLAIEALIGTGNHYHEANCIAEWLHLKGEEAVOLI
           RLSLNMNGDYASALOQGNKLAYPDLEPWALACEYRGLGSALESRLRLARSODPRI
           QTFVGNREQDKT"
           complement(6937. .7200)
gene      /gene="yscf"
CDS       /complement(6937. .7200)
           /gene="yscf"
           /codon_start=1
           /translation="top proteins translocation protein F homolog"
           /protein_id="AAC62549.1"
           /db_xref="GI:2996226"
           /translation="MSNFGFTKGTDIADDAVAQTLKPPDDANKAVNDSIAALKDK
           PDNPALLADLQHSINKMSVIYNISIVSRKDKLMQGLIKFP"
           complement(7201. .7401)
gene      /gene="ysce"
CDS       /complement(7201. .7401)
           /gene="ysce"
           /codon_start=1
           /translation="top proteins translocation protein E homolog"
           /protein_id="AAC62550.1"
           /db_xref="GI:2996227"
           /translation="MTQLEQDLHVEVYRSITQMLEMALTLKDKMRGGDAQOYVA
           QRESKALESAIAIHHYAGDLK"
           complement(7398. .8657)
gene      /gene="yscd"
CDS       /complement(7398. .8657)
           /gene="yscd"
           /codon_start=1
           /translation="top proteins translocation protein D homolog"
           /protein_id="AAC62551.1"
           /db_xref="GI:2996228"
           /translation="MSWYCRFYQGRHNGVEVELPHGRQVPGSDPLQSDIVLSDEIAP
           VHLVAMDEEIRITLDSAEPLQGLPVPITLIRASQCLEVGFILMTFVAVGQPLEE
           TLOVPTQKREKIDLPKSRGLIGIGVLSLILTLFLGLGILGILREYNODQOIEQV
           RRLATIAVYKDVVLTSPKKEGEPMLLGIYIDNARSLQNFLESHGIPFELRSMF
           ELRGAERILORLGYHGEIVSLAPQAGMIGNGEVSSEIQOKRIDSILQAEVPLLQV
           ESKYRIAGNOKRKDALIEQGLDSDTVNVKGLIELRGVQNDKLNSENFLOQTR"
           complement(8657. .9999)

```

Query Match 59.4%; Score 930; DB 2; Length 70504;
 Best Local Similarity 99.9%; Pred. No. 0;


```

LOCUS      AF074612      70559 bp      DNA      circular      BCT      07-APR-2000
DEFINITION Versinia pestis plasmid pCD1, complete plasmid sequence.
ACCESSION  AF074612 M25810
VERSION    AF074612.1 GI:3822037
KEYWORDS
SOURCE     Versinia pestis.
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Versinia pestis.
REFERENCE  1 (bases 43318 to 44840)
            Leung,K.Y. and Straley,S.C.
            The yopM gene of Versinia pestis encodes a released protein having
            homology with the human platelet surface protein GpIIb alpha
            J. Bacteriol. 171 (9), 4623-4632 (1989)
JOURNAL    89359090
MEDLINE    2 (bases 1 to 70559)
REFERENCE  Perry,R.D., Straley,S.C., Featherston,J.D., Rose,D.J., Gregor,J. and
            Blattner,F.R.
            DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1
            of Versinia pestis KIM5
            Infect. Immun. 66 (10), 4611-4623 (1998)
JOURNAL    98427122
MEDLINE    3 (bases 43318 to 44840)
REFERENCE  Straley,S.C.
            Direct Submission
            Submitted (26-APR-1993) Microbiology and Immunology, University of
            Kentucky, MS415 Medical Center, Lexington, KY 40536-0084, USA
            4 (bases 1 to 70559)
            Perry,R.D., Straley,S.C., Featherston,J.D., Rose,D.J., Gregor,J. and
            Blattner,F.R.
            Direct Submission
            Submitted (25-JUN-1998) Microbiology and Immunology, University of
            Kentucky, MS415 Medical Center, Lexington, KY 40536-0084, USA
            On Apr 7, 2000 this sequence version replaced gi:155539.
FEATURES
            source
                1..70559
                /organism="Versinia pestis"
                /plasmid="pCD1"
                /strain="KIM5"
                /db_xref="taxon:632"
                57..368
                /gene="Y0001"
                57..368
                /gene="Y0001"
                /note="O103: 43 pct identical (0 gaps) to 100 residues of
                an approx. 200 aa protein GENEPRF: gi1537126, orf_O198
                Escherichia coli"
                /transl_table=1
                /codon_start=1
                /product="unknown"
                /protein_id="AAC69758.1"
                /db_xref="GI:3822038"
                /translation="WHQSSGGAAGSRILSLIMROSGIVYVWRLARLRMEGLASROP
                GKPYRREREVSLASPDLLKRFQSPBNRWWSGYSIYIVNGMGTALAYIDLYSPH
                W"
                665..1033
                /gene="nuc"
                /note="Y0002"
                665..1033
                /gene="nuc"
                /codon_start=1
                /transl_table=11
                /product="endonuclease"
                /protein_id="AAC69759.1"
                /db_xref="GI:3822039"
                /translation="MDTKLQHTPIGTAVDYPYNTKSGGKRLRCPDVIHYRMDLL
                AVNGIPRYVNSFKALHDKVIYDGKNTQKSGFNSQAAYNSSENTLLIMGDETVOQ
                AYLQYOSRNKKGIDMRSY"
                1171..1425
                /gene="repB"
                /note="Y0003"
                1171..1425
                /gene="repB"

```

```

misc_RNA
    /codon_start-1
    /transl_table-11
    /product-"replication protein"
    /protein_id-"AAC69760.1"
    /db_xref-"GI:3822040"
    /translation-"NSQIENAVTSSSKRAYRKNQNPITGAEKQMSVSRKKEHKAIV
    FIONDKNELDQEDSGLTQTEMERWQREKAATNA"
    complement(1560, 1649)
    /note-"antisense RNA"
    /product-"copa"
    1667..1741
    /gene-"tap"
    /note-"y0004"
    1667..1741
    /gene-"tap"
    /codon_start-1
    /transl_table-11
    /product-"repa translation protein"
    /protein_id-"AAC69761.1"
    /db_xref-"GI:3822041"
    /translation-"MFRKQYLLRLLLPCNISAGRCD"
    1734..2600
    /gene-"repa"
    /note-"y0005"
    1734..2600
    /gene-"repa"
    /codon_start-1
    /transl_table-11
    /product-"replication protein"
    /protein_id-"AAC69762.1"
    /db_xref-"GI:3822042"
    /translation-"MTNQALFTHHYQYKNPNPEPTPREGKTLTPFCRKLAKAEF
    TRPFSEMHVAFARSLSLRMPPLLRRAIDALOGMCFHDPPLANRQSRSTNAI
    ECGLETSKSGNLSITRATRALRLELGYOTEDYDQICNIPDTITPPALESA
    IDVSPVAVAAARSRYEMENQREKQRLRLEMDLIAANRFRPRPSYTERAK
    GIKRARADVDRTDRDIEATVNRQLTRIAEGREFGNIDAVREKARVKERMLMSR
    NNNYTRLATGAT"
    1741..2935
    /note-"oriR"
    complement(3427, 3645)
    /gene-"y0006"
    complement(3427, 3645)
    /gene-"y0006"
    /note-"OriF (f72): 42 pct identical (0 gaps) to 33
    residues of an approx. 216 aa protein GENEPT: g112055297,
    Imp2 Xenopus laevis"
    /codon_start-1
    /transl_table-11
    /product-"unknown"
    /protein_id-"AAC69763.1"
    /db_xref-"GI:3822043"
    /translation-"MRSPVAGCSYTVVYVQLNHDHKSAYKHRLYHVLGQASS
    AQPVSMPHKGKAPADINYSVDNKR"
    4758..5186
    /gene-"y0008"
    4758..5186
    /gene-"y0008"
    /note-"OriF (ol42): 31 pct identical (1 gap) to 48
    residues of an approx. 104 aa protein GENEPT: g112149940,
    Ori1 P. syringae"
    /codon_start-1
    /transl_table-11
    /product-"unknown"
    /protein_id-"AAC69764.1"
    /db_xref-"GI:3822044"
    /translation-"MINTFTLPRKISHFGIDKLSODEYGLCEIILNDRVIMLRA
    DEILNKLTLGLPILGFSGEARSASOLEFCYSINALNKGCFAPWSEELGIAEKHL
    SLDELNENVSSEIANFYDMLSLVSLPATAFSYSYSIG"
    5204..7402
    /gene-"ypka"
    /note-"y0009"
    5204..7402
    /gene-"ypka"

```

```

gene
    /codon_start-1
    /transl_table-11
    /product-"secreted protein kinase"
    /protein_id-"AAC69765.1"
    /db_xref-"GI:3822045"
    /translation-"MKSVMKMGTPPSSISLAKAHERISQHWNPYGEINIGKRYRI
    DNQVLRNPHSGESLFRGVGKIFSGKFNESIANLDTLAAQKTSQELRSDIPN
    ALSNLFQAPQTELPGLGWKGEPLSGAPLEGRVAVETDFAEGSHIS1IETKQRL
    VAKIERISIAEGHLEFALEKAYKHIIYTAGKHPLAVHGMNAVYPGNRKEEALIMDEVD
    GWRCSDTLRLTADSKCKGKINSEAWGRIKFAHRLDLVYVNLAKAGVNDIKRGN
    VEDRASGEVYIDGLHSRSGEQPKGFESRKAPBLGYVNLGASRSKSVFLVYSTLLH
    CIEGFENKPEIKPNQGLRITSEPAHVDENGYPIHRGIAVETAYTRFTDILGVS
    ADSRDSNEARLHEFLSDGTIDEESAKOILADTKAEERGGVDKQKLFNSLITKRYVI
    LIRTLSSAATKQMDMGVLSDDTLMVALKAERGGVDKQKLFNSLITKRYVI
    EDYVKGREGDTKNSSTEVSYPKSNFMLSYEPRIORLOKHLDPHSPSDIGSLVRAH
    KHELTLEVIYVTLISQOGQPVSEETGYFLNLTLEAKITISQOLNTIQOOESKAKDST
    LNRSGSMADVAROSLQRRDSIQPVYKGTEDQYTIHQOMAAHAAITLQEVSEPTD
    MNFTVDSIPLTILQGRSSLMDEHLVEQREKIRELTITTAERLRLERBM"
    7798..8664
    /gene-"yopJ"
    /note-"y0010"
    7798..8664
    /gene-"yopJ"
    /codon_start-1
    /transl_table-11
    /product-"targeted effector protein"
    /protein_id-"AAC69766.1"
    /db_xref-"GI:3822046"
    /translation-"NIGPISQINISGGISEKETSLSJNEBKNIITOLETIDSGSW
    FHNKSRMDVEVPALVIOANNKYPENMLNFTSLDISIEIKNVIENGVSRRPIN
    MEGGIIHFSVIDYKHNGKTSILPEPANTFSGMAMAIIRKTAIEGYVSSRPIIN
    VMEDIORSSSECGIFSLAKKLYIERPSLKHIDNKGILSDENLPHDKIDPYL
    PVTFKHQGGKRLMEYINTNQGYVNNKNEITVAPFNPKNSIVDGKELSVYKH
    KRIAYKILKLYK"
    9858..10064
    /gene-"y0012"
    9858..10064
    /gene-"y0012"
    /note-"O68: 45 pct identical (0 gaps) to 68 residues of an
    approx. 560 aa protein GENEPT: g112162435, hypothetical"
    /codon_start-1
    /transl_table-11
    /product-"unknown"
    /protein_id-"AAC69767.1"
    /db_xref-"GI:3822047"
    /translation-"YAVGRKNLPLGSLRAGORNASIISLETAKLNHDPYWLRYD
    LKLETPNSQNLALPYAENRFS"
    10347..11753
    /gene-"yopH"
    /note-"y0013"
    10347..11753
    /gene-"yopH"
    /codon_start-1
    /transl_table-11

```

Query Match 59.4%; Score 930; DB 2; Length 70559;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 586 ATGATTAGACCTTACGACAAAACCCAAACATTTATTAGAGATCTAGAAAAAGTTAG 645
DB 37919 ATGATTAGACCTTACGACAAAACCCAAACATTTATTAGAGATCTAGAAAAAGTTAG 37918
QY 646 GTGGAAACAATTACTGCTATGGTCTTCAAGTTTGAAGAATGGTTCAGTTAGTCAAA 705
DB 37979 GTGGAAACAATTACTGCTATGGTCTTCAAGTTTGAAGAATGGTTCAGTTAGTCAAA 38038
QY 706 GATAAAATATGATTTTCCATTAAATGATCCAGAAAAGATTCGAGATTTTGGCC 765
DB 38039 GATAAAATATGATTTTCCATTAAATGATCCAGAAAAGATTCGAGATTTTGGCC 38098
QY 766 AATAGAGTAATTACTGATGATTCGATTCAGAAAATCCATGCTATTTCTACCC 825
|||||

```

Db 38099 AATGAGTAACTGATGATATGCAATTCCTCAGAAAATCTAGCTATTTTCTACCC 38158
 QY 826 GAGGATACCATCTCTTAAGCGGTCATTTATGACAAACCAATGCAAAATGGCATCAAGCA 885
 Db 38159 GAGGATGCAATCTTAAAGCGGTCATTTATGACAAACCAATGCAAAATGGCATCAAGCA 38218
 QY 886 GTAAAAAGATCTCTGATGATGATGCGGAAATACAAATGGGAATGGCGGTCATGCGCA 945
 Db 38219 GTAAAAAGATCTCTGATGATGATGCGGAAATACAAATGGGAATGGCGGTCATGCGCA 38278
 QY 946 GTAAATGATCTCTCTTAAAGCGGTCATTTATGATGATGATTTTGAAGATGTT 1005
 Db 38279 GTAAATGATCTCTCTTAAAGCGGTCATTTATGATGATGATTTTGAAGATGTT 38338
 QY 1006 GATTCATGATCATGATGATGATGCGGTCATGATGATGATGATGATGATGATGATGAT 1065
 Db 38339 GATTCATGATCATGATGATGATGCGGTCATGATGATGATGATGATGATGATGATGAT 38398
 QY 1066 ACCGCGGAATTAAGATTTATTCAGTTATTCAGCGGAATTAATAGCATCTGTCTAGT 1125
 Db 38399 ACCGCGGAATTAAGATTTATTCAGTTATTCAGCGGAATTAATAGCATCTGTCTAGT 38458
 QY 1126 ATGCGGACCATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
 Db 38459 ATGCGGACCATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 38518
 QY 1186 TATACAGATGAAGATTTTAAAGCGGTCATGATGATGATGATGATGATGATGATGATGAT 1245
 Db 38519 TATACAGATGAAGATTTTAAAGCGGTCATGATGATGATGATGATGATGATGATGATGAT 38578
 QY 1246 CAACCCACCATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
 Db 38579 CAACCCACCATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 38638
 QY 1306 GGAAGTGAAGATTTTAAAGCGGTCATGATGATGATGATGATGATGATGATGATGATGAT 1365
 Db 38639 GGAAGTGAAGATTTTAAAGCGGTCATGATGATGATGATGATGATGATGATGATGATGAT 38698
 QY 1366 AAGATTAATGAATTTATTCAGTTATTCAGCGGAATTAATAGCATCTGTCTAGT 1425
 Db 38699 AAGATTAATGAATTTATTCAGTTATTCAGCGGAATTAATAGCATCTGTCTAGT 38758
 QY 1426 AAGCATGCTGTAGCGCAAAACCACTGATGATGATGATGATGATGATGATGATGATGAT 1485
 Db 38759 AAGCATGCTGTAGCGCAAAACCACTGATGATGATGATGATGATGATGATGATGATGAT 38818
 QY 1486 GCTATTGAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1545
 Db 38819 GCTATTGAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 38878
 QY 1546 GATGACAGCTGTGTAATGA 1566
 Db 38879 GATGACAGCTGTGTAATGA 38899

RESULT 5
 A46411
 LOCUS A46411 1014 bp DNA PAT 07-MAR-1997
 DEFINITION Sequence 1 from Patent WO9524475.
 ACCESSION A46411
 VERSION A46411.1 GI:2300612
 KEYWORDS
 SOURCE
 ORGANISM
 Yersinia pestis.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.
 REFERENCE
 1 (bases 1 to 1014)
 AUTHORS Titball, R.W., Williamson, E.D. and Leary, S.E.
 TITLE VACCINE COMPOSITIONS
 JOURNAL Patent: WO 9524475-A 1 14-SEP-1995;
 COMMENT
 FEATURES
 Other publication AU 1853995 950925.
 Location/Qualifiers

source 1..1014
 /organism="Yersinia pestis"
 /db_xref="taxon:632"
 CDS 1..990
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA02893.1"
 /db_xref="GI:2300613"
 /translation="MSERIRAEONPQHFIEDLEKRVLEQTLGHSSTLEELVQLVKD
 KNIDISIDPRKDSVFNANVTIDIDLELKLILATFPEPAILGGHYDNOLOGIK
 RVKELESSPTQWELRAFMVMEISLADRIDDIIVYDSNMHNDASKREEL
 AELRAELKIVYSIQAEIKHLSGGTINIHDKSLMDKNLYGYDEEILFASAEYKI
 LEKPOOTTIOYDSEKRIYSIKDEFSNKRGAAGNKGNSYSYKDNELSHFATTC
 SDKSRPLNDIVSOKTTOJSDITSRNSAIEALNRIQKIVDSVMRLDDTSK"
 BASE COUNT 346 a 181 c 201 g 286 t
 ORIGIN
 Query Match 59.2%; Score 927; DB 5; Length 1014;
 Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;
 Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 589 ATAGAGCCTAGCAACAAAACCAACATTTTATGAGATCTAGAAAAGTTAGGTG 648
 Db 13 ATAGAGCCTAGCAACAAAACCAACATTTTATGAGATCTAGAAAAGTTAGGTG 72
 QY 649 GAACAACCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
 Db 73 GAACAACCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 132
 QY 709 AAAAATATGATTTTCCATTAATATGATGATGATGATGATGATGATGATGATGATGAT 768
 Db 133 AAAAATATGATTTTCCATTAATATGATGATGATGATGATGATGATGATGATGATGAT 192
 QY 769 AGAGTATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828
 Db 193 AGAGTATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252
 QY 829 GATACATTTCTTAAAGCGGTCATTTATGACAAACCAATGCAATGCAACGAGTA 888
 Db 253 GATACATTTCTTAAAGCGGTCATTTATGACAAACCAATGCAATGCAACGAGTA 312
 QY 889 AAGAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 Db 313 AAGAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
 QY 949 ATGCATTTCTTTAAAGCGGTCATGATGATGATGATGATGATGATGATGATGATGAT 1008
 Db 373 ATGCATTTCTTTAAAGCGGTCATGATGATGATGATGATGATGATGATGATGATGAT 432
 QY 1009 TCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
 Db 433 TCAATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
 QY 1069 GCCGATTAATGATTTATTCAGTTATTCAGCGGAATTAATAGCATCTGTCTAGT 1128
 Db 493 GCCGATTAATGATTTATTCAGTTATTCAGCGGAATTAATAGCATCTGTCTAGT 552
 QY 1129 GCCGATTAATGATTTATTCAGTTATTCAGCGGAATTAATAGCATCTGTCTAGT 1188
 Db 553 GCCGATTAATGATTTATTCAGTTATTCAGCGGAATTAATAGCATCTGTCTAGT 612
 QY 1189 ACAGATGAAGATTTTAAAGCGGTCAGATGATGATGATGATGATGATGATGATGATGAT 1248
 Db 613 ACAGATGAAGATTTTAAAGCGGTCAGATGATGATGATGATGATGATGATGATGATGAT 672
 QY 1249 ACCACATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
 Db 673 ACCACATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
 QY 1309 ACTGAGATTAAGAAACCGGCGGTCGTTGGTATCTGAAAATCTAGCTTATATATA 1368

Db 733 AGTGAATATAAAGACACCGGGCGTTGGTATCTGAAAACTCATCTTATATATAA 792

QY 1369 GATATATATGATATATCTCATCTTGGCCACACCTGCTCGATTAAGTCCAGGCCCTCAAC 1428

Db 793 GATATATATGATATATCTCATCTTGGCCACACCTGCTCGATTAAGTCCAGGCCCTCAAC 852

QY 1429 GACTGGTGGCCAAAAACACACCTGCTGCTATATATATATATATATATATATATAT 1488

Db 853 GACTGGTGGCCAAAAACACACCTGCTGCTATATATATATATATATATATATATATATAT 912

QY 1489 ATTGAACACTGAAACCGTTTCATGAGAAATATGATTCAGTATGCAACGCTGCTAGAT 1548

Db 913 ATTGAACACTGAAACCGTTTCATGAGAAATATGATTCAGTATGCAACGCTGCTAGAT 972

QY 1549 GACACGCTGCTGTAATGA 1566

Db 973 GACACGCTGCTGTAATGA 990

RESULT 6

LOCUS A56793 1014 bp DNA PAT 03-MAR-1998

DEFINITION Sequence 1 from Patent WO9628551.

ACCESSION A56793

VERSION A56793.1 GI:3712808

KEYWORDS

SOURCE Yersinia pestis.

ORGANISM Yersinia pestis

REFERENCE 1 (bases 1 to 1014)

AUTHORS Tibbitt, R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C., Bennett and Alice, M.

TITLE VACCINES FOR PLAGUE

JOURNAL Patent: WO 9628551-A 1 19-SEP-1996;

COMMENT SECR DEFENCE (GB)

Other publication ZA 9602036 960716

Other publication AU 4951196 961002.

FEATURES

source

1. 1014

/organism="Yersinia pestis"

/db_xref="taxon:632"

<1.990

/note="unnamed protein product"

/codon_start=1

/transl_table=1

/protein_id="CA03419.1"

/db_xref="GI:4530021"

/translation="ISEFTIRAYEQNPQHIEDLEKRVEDLTGHSSVLEELYQVLYKD

KNDISKYDPRKDSVFANRVITDDIELKTLAYFLPEDALDKGHDNOMQNGIK

RKVEFLSSPNQWELRAFMVNHFSITDRIDDILKIVSMNHGHRASRLREIL

AELTALIKTIVYQAEINKHSSGFINIDKISINMDKLYGTDEEIFKASAEYKI

LEMPQTTIOVDSSEKKIYSINDFLGSEKMRGALINIKNSYVYNNDNNEISFATTC

SDKSRPLNDLVSKTQLOSDIRSPNSALFNRFQIKIDSVQRLDIDTSGR"

BASE COUNT 346 a 181 c 201 g 286 t

ORIGIN

Query Match 59.2%; Score 927; DB 5; Length 1014;

Best local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 589 ATTAGAGCCCTACGAAACAAACCAACATTTTATGAGATCTAGAAAAAGTTAGGCTG 648

Db 13 ATTAGAGCCCTACGAAACAAACCAACATTTTATGAGATCTAGAAAAAGTTAGGCTG 72

QY 649 GAACCACTTCTGCTGCTGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 708

Db 73 GAACCACTTCTGCTGCTGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 132

QY 709 AAAAATATAGATATTTTCCATTAATATGATCCGAAAGATTCGGAGTTTGGCCAAAT 768

Db 133 AAAAATATAGATATTTTCCATTAATATGATCCGAAAGATTCGGAGTTTGGCCAAAT 192

QY 769 AGATATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828

Db 193 AGATATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252

QY 829 GATATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888

Db 253 GATATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312

QY 889 AAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 948

Db 313 AAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 372

QY 949 ATGATATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008

Db 373 ATGATATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432

QY 1009 TCAATGATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068

Db 433 TCAATGATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492

QY 1069 GCGGAATTAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128

Db 493 GCGGAATTAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552

QY 1129 GCGACATTAATATATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188

Db 553 GCGACATTAATATATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 612

QY 1189 ACAGATTAAGATTTTAAAGCCAGGAGTACAAATTTCTCGAAGAAATGCTCAA 1248

Db 613 ACAGATTAAGATTTTAAAGCCAGGAGTACAAATTTCTCGAAGAAATGCTCAA 672

QY 1249 ACCGACATGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308

Db 673 ACCGACATGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732

QY 1309 AGTGAATTAAGAAACCGGGGCTGGTATCTGAAAACTCATCTTATATATAA 1368

Db 733 AGTGAATTAAGAAACCGGGGCTGGTATCTGAAAACTCATCTTATATATAA 792

QY 1369 GATATATATGATATATCTCATCTTGGCCACACCTGCTCGATTAAGTCCAGGCCCTCAAC 1428

Db 793 GATATATATGATATATCTCATCTTGGCCACACCTGCTCGATTAAGTCCAGGCCCTCAAC 852

QY 1429 GACTGGTGGCCAAAAACACACCTGCTGCTATATATATATATATATATATATATATAT 1488

Db 853 GACTGGTGGCCAAAAACACACCTGCTGCTATATATATATATATATATATATATATATAT 912

QY 1489 ATTGAACACTGAAACCGTTTCATGAGAAATATGATTCAGTATGCAACGCTGCTAGAT 1548

Db 913 ATTGAACACTGAAACCGTTTCATGAGAAATATGATTCAGTATGCAACGCTGCTAGAT 972

QY 1549 GACACGCTGCTGTAATGA 1566

Db 973 GACACGCTGCTGTAATGA 990

RESULT 7

LOCUS A56808 1462 bp DNA PAT 03-MAR-1998

DEFINITION Sequence 16 from Patent WO9628551.

ACCESSION A56808

VERSION A56808.1 GI:3712821

KEYWORDS

SOURCE Yersinia pestis.

ORGANISM Yersinia pestis

REFERENCE 1 (bases 1 to 1462)

AUTHORS Tibbitt, R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C., Bennett and Alice, M.

TITLE VACCINES FOR PLAGUE

JOURNAL

Patent: WO 9628551-A 16 19-SEP-1996;

SECURE DEFENCE (GB)

Other publication ZA 9602036 960716
Other publication AU 4951196 961002.

FEATURES Location/Qualifiers

SOURCE

1. 1462

/organism="Yersinia pestis"

/db_xref="taxon:632"

<8. 1450

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CAA03421.1"

/db_xref="GI:3712822"

/translation="ADLSTSTATITLVEPARITLYREGAPITINDGNIDTELVG
TLVIGRTGTSVNTFDAGDPMTLFTSDGNHOFITKVGKDSRUPDISPKV
NGENLVGDVVLATGSDPEFVRSIGSKGLAAGYATVATVYSNOSIGRIAYE
ONPOHIEDLEKRVLEOLGHSVLEELVOLKRNIDISIKYPRDSSEFARVI
TDIELEKRIAYEPEDAIKLGHHYDNLONGIRKREPLESSSPNOMELRAFAVM
HESLADRIDDDILKVIYDSNHHGDASKIREELAEIKYISYQAEIKHLS
SGTINHRKSNINDKNIYGTDEIEFASAEKYLEMPOTITVOSEKIVSIKD
FLGSENRKRTGALGNLKNISYINKDNELSHFATTCSDKSRPLNDLVSKTTQLSDITS
RNSAIEALNFIKRYDSVMORLDDTSGK"

BASE COUNT 476 a 286 c 300 g 400 t

ORIGIN

Query Match 59.2%; Score 927; DB 5; Length 1462;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 589 ATTACAGCTAGACAAACCCACACATTTATTGAGATCTAGAAAAGTAGGGG 648
 DB 473 ATTACAGCTAGACAAACCCACACATTTATTGAGATCTAGAAAAGTAGGGG 532
 QY 649 GAACAACCTAGCTAGCTGCTCAGTTTGAAGAATGGTTCAGTTCAGTAAGAT 708
 DB 533 GAACAACCTAGCTAGCTGCTCAGTTTGAAGAATGGTTCAGTTCAGTAAGAT 592
 QY 709 AAAAAATATGATATTTCCATTAATATGATCCAGAAAAGTTCGAGGTTTTCGCAAT 768
 DB 593 AAAAAATATGATATTTCCATTAATATGATCCAGAAAAGTTCGAGGTTTTCGCAAT 652
 QY 769 AGAGATATCTAGATGATTCGAATTCGTCAGAAAATCTGATCTATTTTCAACCGAG 828
 DB 653 AGAGATATCTAGATGATTCGAATTCGTCAGAAAATCTGATCTATTTTCAACCGAG 712
 QY 829 GATACCATCTTAAAGCGCGTATATGACAAACCTGCAAAATGGCATCAAGCGAGTA 888
 DB 713 GATACCATCTTAAAGCGCGTATATGACAAACCTGCAAAATGGCATCAAGCGAGTA 772
 QY 889 AAGAGATCTTGAATCATGCGCGAATACACATGGAATGGGCGCTTCATGCGAGTA 948
 DB 773 AAGAGATCTTGAATCATGCGCGAATACACATGGAATGGGCGCTTCATGCGAGTA 832
 QY 949 ATGCATTTCTCTTAAACCGCGATGATGATGATATTTTGAAGTATGTTGAT 1008
 DB 833 ATGCATTTCTCTTAAACCGCGATGATGATGATATTTTGAAGTATGTTGAT 892
 QY 1009 TCAATGAATCATGATGATGCGCGTACGAAGTTCGTAAGAAATAGCTGAGCTTAC 1068
 DB 893 TCAATGAATCATGATGATGCGCGTACGAAGTTCGTAAGAAATAGCTGAGCTTAC 952
 QY 1069 GCCGATTTAAAGATTTATCACTTATTCAGCGGAATTAATAGCATCTGCTAGTACT 1128
 DB 953 GCCGATTTAAAGATTTATCACTTATTCAGCGGAATTAATAGCATCTGCTAGTACT 1012
 QY 1129 GCCGATTTAAATCATATATATCATATTAATTCATGATTAATTAATGTTAT 1188
 DB 1013 GCCGATTTAAATCATATATATCATATTAATTCATGATTAATTAATGTTAT 1072
 QY 1189 ACAGATGAAGATTTTAAACCGCGAGATACAAAATCTGAGAAATGCTCTCA 1248

DB 1073 ACAGATGAAGATTTTAAACCGCGAGATACAAAATCTGAGAAATGCTCTCA 1132
 QY 1249 ACCACATTCAGCTGATGAGGAGCGAGAAAATATGCTCATTAAGACTTTCTTGA 1308
 DB 1133 ACCACATTCAGCTGATGAGGAGCGAGAAAATATGCTCATTAAGACTTTCTTGA 1192
 QY 1309 AGTGAATTAAGAAACCGGCGTGGGTATCTGAAAATCTCATCTATATATTA 1368
 DB 1193 AGTGAATTAAGAAACCGGCGTGGGTATCTGAAAATCTCATCTATATATTA 1252
 QY 1369 GATTAATTAAGATTTATCTCATCTTTCGACACCTGCTGATTAAGTCCAGCGCTCAAC 1428
 DB 1253 GATTAATTAAGATTTATCTCATCTTTCGACACCTGCTGATTAAGTCCAGCGCTCAAC 1312
 QY 1429 GACTGGTTAGCCAAAACCACTAGCTGCTGATTAATTAATTAATTAATTA 1488
 DB 1313 GACTGGTTAGCCAAAACCACTAGCTGCTGATTAATTAATTAATTAATTA 1372
 QY 1489 ATTGAAGCCTGACCGTTTCATGAAATATGATTCAGTATGCAACGCTGTAGAT 1548
 DB 1373 ATTGAAGCCTGACCGTTTCATGAAATATGATTCAGTATGCAACGCTGTAGAT 1432
 QY 1549 GACAGCTCTGTAATGA 1566
 DB 1433 GACAGCTCTGTAATGA 1450

RESULT 8

A56814

Sequence 22 from Patent WO9628551.

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

CDS

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

ORIGIN

BASE COUNT

QY	589	ATTAGAGCCTACGAGCAAAACCCAAACATTTTATTTGAGATCTGAGAAAAGTTAGGGTG	648
Db	541	ATTAGAGCCTACGAGCAAAACCCAAACATTTTATTTGAGATCTGAGAAAAGTTAGGGTG	600
QY	649	GAAACAACCTTACTGTCATCGTTCTTCACTTTTATAGAAATTTGGTTCAGTTAGTCAAGAT	708
Db	601	GAAACAACCTTACTGTCATCGTTCTTCACTTTTATAGAAATTTGGTTCAGTTAGTCAAGAT	660
QY	709	AAATAATAGATTTTTCATTAATTAATGATATCCGAAAAAGTTTGGAGGTTTGGCAAT	768
Db	661	AAAAATATAGATTTTTCATTAATTAATGATATCCGAAAAAGTTTGGAGGTTTGGCAAT	720
QY	769	AGAGTAATTTACTGATGATATGCAATTTGCTCAAGAAAATCCTAGCTTATTTTCCACCGAG	828
Db	721	AGAGTAATTTACTGATGATATGCAATTTGCTCAAGAAAATCCTAGCTTATTTTCCACCGAG	780
QY	829	GATACCANTCTTAAAGCGGCTCATTTATGACACCAACCTGCACAAAATGGCATCAACGAGTA	888
Db	781	GATGCCATTTCTTAAAGCGGCTCATTTATGACACCAACCTGCACAAAATGGCATCAACGAGTA	840
QY	889	AAAGAGTTCCTTGATCATATCCGCCAATACAAATGGGAATTTGGGGGCTCATAGCGAGTA	948
Db	841	AAAGAGTTCCTTGATCATATCCGCCAATACAAATGGGAATTTGGGGGCTCATAGCGAGTA	900
QY	949	ATGCATTTCTCTTTAAACCGCGATCGTATCATGATATATTTTGAAGAATGATTTGGAT	1008
Db	901	ATGCATTTCTCTTTAAACCGCGCGATCGTATCATGATATATTTTGAAGAATGATTTGGAT	960
QY	1009	TCAATGAATCATCATGTTGATGCCCGTAGCAAGTTGCCGTGAGAAATTTAGCTGAGCTTACC	1066
Db	961	TCAATGAATCATCATGTTGATGCCCGTAGCAAGTTGCCGTGAGAAATTTAGCTGAGCTTACC	1022
QY	1069	GCCGAATTTAAAGATTTATTCAGTTATTCACAGCCGAATTTAATAGCATCTGCTACTAGT	1122
Db	1021	GCCGAATTTAAAGATTTATTCAGTTATTCACAGCCGAATTTAATAGCATCTGCTACTAGT	1080
QY	1129	GCGACCATTAATATTCATGATTAATTCATTAATCTCATGATATAAAATTTATATAGTTAT	1188
Db	1081	GCGACCATTAATATTCATGATTAATTCATTAATCTCATGATATAAAATTTATATAGTTAT	1144
QY	1189	ACAGATGGAAGATTTTTTAAAGCCAGGCGAGATGACAAATTTCTCGAAGAAATGCTCTAA	1244
Db	1141	ACAGATGGAAGATTTTTTAAAGCCAGGCGAGATGACAAATTTCTCGAAGAAATGCTCTAA	1200
QY	1249	ACCACATTCAGTGATGATGGAGCGAGAAAAAATTAATCTCGATAAGAAAGACTTCTGGG	1308
Db	1201	ACCACATTCAGTGATGATGGAGCGAGAAAAAATTAATCTCGATAAGAAAGACTTCTGGG	1264
QY	1309	AGTGAGAAATTAAGAACCCGGGGCTTGGGTATCTGAAAAACCTCATACTTTAATATAA	1366
Db	1261	AGTGAGAAATTAAGAACCCGGGGCTTGGGTATCTGAAAAACCTCATACTTTAATATAA	1322
QY	1369	GATATATATGAATTAATCTCATATTTGGCACACCACTGCTCGGATATAAGTCCAGCGGCTCAAC	1422
Db	1321	GATATATATGAATTAATCTCATATTTGGCACACCACTGCTCGGATATAAGTCCAGCGGCTCAAC	1380
QY	1429	GACTGTTGTAGCCAAAAAACACACTCAGCTGCTGATATTTACATCAGATTTAATTAATAGCT	1488
Db	1381	GACTGTTGTAGCCAAAAAACACACTCAGCTGCTGATATTTACATCAGATTTAATTAATAGCT	1444
QY	1489	ATTGAGACACTGAAACCGTTTCATTCAGAAATATGATTCAGTGTATGCAACGCTGCTAGAT	1544
Db	1441	ATTGAGACACTGAAACCGTTTCATTCAGAAATATGATTCAGTGTATGCAACGCTGCTAGAT	1500
QY	1549	GACACGCTGCTGTAATATGA 1566	
Db	1501	GACACGCTGCTGTAATATGA 1518	

RESULT	9
A46413	
LOCUS	A46413
DEFINITION	Sequence 3 from Patent WO9524475.
	1014 bp DNA PAT
	07-MAR-1997

ACCESSION	A46413
VERSION	A46413.1
KEYWORDS	GI:2300614
SOURCE	Yersinia pestis.
ORGANISM	Yersinia pestis Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
REFERENCE	1 (bases 1 to 1014)
AUTHORS	Tibball,R.W., Williamson,E.D. and Leary,S.E.
TITLE	VACCINE COMPOSITIONS
JOURNAL	Patent: WO 9524475-A 3 14-SEP-1995; SECR DEFENCE BRIT (GB) Other publication AU 1853995 950925. Location/Qualifiers
COMMENT	1..1014
FEATURES	/organism="Yersinia pestis" /db_xref="taxon:632"
source	1..987
CDS	/note="unnamed protein product; Protein sequence is in conflict with the conceptual translation" /codon_start=1 /transl_table=1 /protein_id="CAO2894.1" /db_xref="gi:2300615" /translation="GIPETRAVEQNPOHFIEDLERVYEOLTGHSVLEELVOLYKIKNIIDISIRYDKDESEVANFNVITDDIELLKKIILAYEPEDAILKGHYDQNLNGIRKIFLESPTOMELRAFMAVMHFSLELDLIDLILKYIDSMNHGDASKREELELAETKEITVIOAEIKHLSSSGTINIHKSJNLMOKNLNYSTNNKDELSPASVEYKLEKPDTOIVDSESKRIVTSIKDPGLGSNKRTGALGNKNYSYNKDNELSHRATTCSDKRPNLDLVSRITQTLOSDITSFRNSATELNRIKTDIVSWORLDVISGR"
BASE COUNT	343 a 184 c 205 g 282 t
ORIGIN	
Query Match	58.9% Score 923; DB 5; Length 1014;
Best Local Similarity	99.9%; Pred. No.0's
Matches 973; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	593 GAGCCTAGCAACAACCCACACACTTTTATGAGAGCTAGAAAAAGTTAGGTGGAAC 652
DB	17 GAGCCTAGCAACAACCCACACACTTTTATGAGAGCTAGAAAAAGTTAGGTGGAAC 76
OY	653 AACTACTGTCATGGTCTTCAGTTTTAGAAGATTGGTTCAGTAGTCAAAGATAAAA 712
DB	77 AACTACTGTCATGGTCTTCAGTTTTAGAAGATTGGTTCAGTAGTCAAAGATAAAA 136
OY	713 AATATGATATTTCCATTAAATATGATCCCAGAAAAAGATTGGAGGTTTTGCCATTAG 772
DB	137 AATATGATATTTCCATTAAATATGATCCCAGAAAAAGATTGGAGGTTTTGCCATTAG 196
OY	773 TAATACTGATGATATCGAATGTGCACAGAAAATCCTAGCTTAATTTGACCCGAGATA 832
DB	197 TAATACTGATGATATCGAATGTGCACAGAAAATCCTAGCTTAATTTGACCCGAGATG 256
OY	833 CCATTCTTAAAGCGCGCATATATGCAACAACACTGCCAAATATGGCATCAAGCAGTAAAG 892
DB	257 CCATTCTTAAAGCGCGCATATATGCAACAACACTGCCAAATATGGCATCAAGCAGTAAAG 316
OY	893 AGTCTCTGAATCATCGCCGAATACACAATGGGAATTGCGGGCGTTCAATGGCAGTAATGC 952
DB	317 AGTCTCTGAATCATCGCCGAATACACAATGGGAATTGCGGGCGTTCAATGGCAGTAATGC 376
OY	953 ATTCTCTTTAAGCGCGCATGATGCATGATGATTTTGAAGTGAATGTTGTGATTCAA 1012
DB	377 ATTCTCTTTAAGCGCGCATGATGCATGATGATTTTGAAGTGAATGTTGTGATTCAA 436
OY	1013 TGAATCATCATGATGATGCCGTAGCAGTAGTGGCGTGAAGAAATTAGCTAGCTTACGGCG 1072
DB	437 TGAATCATCATGATGATGCCGTAGCAGTAGTGGCGTGAAGAAATTAGCTAGCTTACGGCG 496
OY	1073 AATTAAAGATTTATTCAGTTATTCAGGCCGAATTAATTAAGCATCTGCTAGTAGTGCA 1132
DB	497 AATTAAAGATTTATTCAGTTATTCAGGCCGAATTAATTAAGCATCTGCTAGTAGTGCA 556

QY	1133	CCATTAATATCCATGATTAATCCATTAATCTCATGATGATAAAATTTATATGTTATACAG	1132
Db	557	CCATTAATATCCATGATTAATCCATTAATCTCATGATGATAAAATTTATATGTTATACAG	616
QY	1193	ATGAAGACATTTTAAAGCCAGCGCAGAGTACAAATTCCTGAGAAAAATGCCTCMAACA	1252
Db	617	ATGAAGACATTTTAAAGCCAGCGCAGAGTACAAATTCCTGAGAAAAATGCCTCMAACA	676
QY	1253	CCATTACAGTGGATGGGAGCGAGAAAAAATATAGTCTGATTAAGACCTTTCTTGGAAATG	1312
Db	677	CCATTACAGTGGATGGGAGCGAGAAAAAATATAGTCTGATTAAGACCTTTCTTGGAAATG	736
QY	1313	AGAAATAAAGAACCGGGGCGCTGGTGGTATCTGAAAAATCTACTATATATTAAGATA	1372
Db	737	AGAAATAAAGAACCGGGGCGCTGGTGGTATCTGAAAAATCTACTATATATTAAGATA	796
QY	1373	ATAATGAATTAATCTACTTTTCCACACACCTCTCGGATAGTCCAGGCCGCTCAACGACT	1432
Db	797	ATAATGAATTAATCTACTTTTCCACACACCTCTCGGATAGTCCAGGCCGCTCAACGACT	856
QY	1433	TGGTTAGCCAAAAACAACACTCAGCTGTGTGATTTATACATCACGTTTATATTCAGTATG	1492
Db	857	TGGTTAGCCAAAAACAACACTCAGCTGTGTGATTTATACATCACGTTTATATTCAGTATG	916
QY	1493	AAGCCTAACCGGTTTCATTAAGAAATATGATTCGATATGCAACGCTCGTAGATGACA	1552
Db	917	AAGCCTAACCGGTTTCATTAAGAAATATGATTCGATATGCAACGCTCGTAGATGACA	976
QY	1553	CGCTGTGTAATGA	1566
Db	977	CGCTGTGTAATGA	990
RESULT	10		
LOCUS	A56795	DNA	
DEFINITION	Sequence 3 from Patent WO9628551.	PAT	03-MAR-1998
ACCESSION	A56795		
VERSION	A56795.1		
KEYWORDS	GI:3712810		
SOURCE	Yersinia pestis.		
ORGANISM	Yersinia pestis.		
REFERENCE	1 (bases 1 to 1014)		
AUTHORS	TiDhall,R.W., Williamson,E.D., Ieary, Sophie,E.C., Oyston,P.C., Bennett and Alice,M.		
TITLE	VACCINES FOR PLAGUE.		
JOURNAL	Patent: WO 9628551-A 3 19-SEP-1996;		
COMMENT	SECR DEFENCE (GB) 2A 9602036 960716		
FEATURES	Other publication AU 4951196 961002.		
SOURCE	Location/Qualifiers		
	1..1014		
	/organism="Yersinia pestis"		
	/db_xref="taxon:632"		
	<1..990		
	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CA03420.1"		
	/db_xref="GI:3712811"		
	/translation="GIPETRAYEDNPQHFIDLEKRYEOLTHGSSVLELVQYKD		
	KNIIDISIKYDRKDSVFANVITDDIELKLILAYFLPEDAILKGHYNOLONGLIK		
	RVKEFLSEPTQWELRAFMVAMFSLADRIDDDILKIVDSNNHGGDARSKEBEL		
	ALIELAEKTIQVIOAEIKHLSSGGINIDKRSINIMDKNLNYGTDEIFKASLEYK		
	LEKHPQITQVDSSEKRIKDFGSENRKRTGALGNKNSYVKNKNNELSHPATTC		
	SDKRPLNDLVISQRTQSLDITSRFNSAIEALNFIQKIDYVMRLDDITSGK"		
BASE COUNT	343 a	184 c	205 g 282 t
ORIGIN			

	Query Match	58.9%;	Score 923;	DB 5;	Length 1014;	
	Best Local Similarity	99.9%;	Pred. No. 0;			
	Matches	973;	Conservative	0;	Mismatches	1;
					Indels	0;
					Gaps	0;
QY	593	GAGCCTACGAACAAACCACACACATTTTATTAGAGATCTAGAAAAAGTTAGGGTGAAC	652			
Db	17	GAGCCTACGAACAAACCACACACATTTTATTAGAGATCTAGAAAAAGTTAGGGTGAAC	76			
QY	653	AACCTACGCGTCATGCTCTTCATGTTTAGAAGATTGGTTAGTTAGTCAAGATAAAA	712			
Db	77	AACCTACGCGTCATGCTCTTCATGTTTAGAAGATTGGTTAGTTAGTCAAGATAAAA	136			
QY	713	ATAATGATATTTCCATTAATATGATCCGAGAAAAGATTGGAGGTTTGGCAATAGAG	772			
Db	137	ATAATGATATTTCCATTAATATGATCCGAGAAAAGATTGGAGGTTTGGCAATAGAG	196			
QY	773	TAATACTAGATGATTCGAATTCGTCAGAAAATCCTAGCTTATTTTCCACCGAGATA	832			
Db	197	TAATACTAGATGATTCGAATTCGTCAGAAAATCCTAGCTTATTTTCCACCGAGATG	256			
QY	833	CCATTCTTAAAGCGCGCTATTATGACAAACCACTGCAAAATGGCATCAAGCAGTAAAG	892			
Db	257	CCATTCTTAAAGCGCGCTATTATGACAAACCACTGCAAAATGGCATCAAGCAGTAAAG	316			
QY	893	AGTTCCTTGATCATCGCCGAATACACAATGGGAATTGGGGCGTTCATGCGAGTATGC	952			
Db	317	AGTTCCTTGATCATCGCCGAATACACAATGGGAATTGGGGCGTTCATGCGAGTATGC	376			
QY	953	ATTTCCTTTAACCGCGCGATGCTATGCATGATGATATTTTGAAGGATGGTGGATTCA	1012			
Db	377	ATTTCCTTTAACCGCGCGATGCTATGCATGATGATATTTTGAAGGATGGTGGATTCA	436			
QY	1013	TGAATCATCATGCTGATGCGCGGTAGCAGATTGCGTGAAGAAATTAGCTGACCTACCGCG	1072			
Db	437	TGAATCATCATGCTGATGCGCGGTAGCAGATTGCGTGAAGAAATTAGCTGACCTACCGCG	496			
QY	1073	AATTAAGATTTTATTCAGTTATTCAGCCGGAATTAATTAAGCATCTGCTAGTAGTGCA	1132			
Db	497	AATTAAGATTTTATTCAGTTATTCAGCCGGAATTAATTAAGCATCTGCTAGTAGTGCA	556			
QY	1133	CCATAAATTCATGATTAATTAATCCATTATTCATGATTAATTAATTAATTAATTAATG	1192			
Db	557	CCATAAATTCATGATTAATTAATCCATTATTCATGATTAATTAATTAATTAATTAATG	616			
QY	1193	ATGAAGAGATTTTAAAGCAGCGCAGAGTACAAAATTCGACGAAAATAGCCCTCAACCA	1252			
Db	617	ATGAAGAGATTTTAAAGCAGCGCAGAGTACAAAATTCGACGAAAATAGCCCTCAACCA	676			
QY	1253	CCATTACAGTGATGGAGCGAGAAAAAATAGTCTCGATTAAGAGACTTCTTGGAAAGTG	1312			
Db	677	CCATTACAGTGATGGAGCGAGAAAAAATAGTCTCGATTAAGAGACTTCTTGGAAAGTG	736			
QY	1313	AGAATTAAGAAACCGGGGCGTTGGGTATCTGAAAAAATCACTACTCTTATTAATAAAGAT	1372			
Db	737	AGAATTAAGAAACCGGGGCGTTGGGTATCTGAAAAAATCACTACTCTTATTAATAAAGAT	796			
QY	1373	ATAATGAATTAATCACTCTTGCCACACACCTGCTCGATTAAGTCCAGCGCGCTCAAGACT	1432			
Db	797	ATAATGAATTAATCACTCTTGCCACACACCTGCTCGATTAAGTCCAGCGCGCTCAAGACT	856			
QY	1433	TGGTTAGCCAAAACCACTCAGCTGCTCGATTAATTCATCAGTTTAAATTCAGCTATTG	1492			
Db	857	TGGTTAGCCAAAACCACTCAGCTGCTCGATTAATTCATCAGTTTAAATTCAGCTATTG	916			
QY	1493	AAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGCTAGATGACA	1552			
Db	917	AAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGCTAGATGACA	976			
QY	1553	CGTCTGCTTAATGA 1566				
Db	977	CGTCTGCTTAATGA 990				


```

RESULT 11
LOCUS AF167309 975 bp DNA BCT 17-APR-2000
DEFINITION Yersinia pestis strain Pestoides F V antigen (lcrv) gene, complete cds.
ACCESSION AF167309
VERSION AF167309.1 GI:7578512
KEYWORDS Yersinia pestis.
SOURCE Yersinia pestis.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
REFERENCE 1 (bases 1 to 975)
AUTHORS Adair, D.M., Worsham, P.L., Hill, K.K., Klevytska, A.M., Jackson, P.J., Friedlander, A.M. and Keim, P.
Diversity in a variable-number tandem repeat from Yersinia pestis J. Clin. Microbiol. 38 (4), 1516-1519 (2000)
20211685
JOURNAL MEDLINE
REFERENCE 2 (bases 1 to 975)
AUTHORS Hill, K.K. and Jackson, P.J.
Diversity in a variable-number tandem repeat from Yersinia pestis J. Clin. Microbiol. 38 (4), 1516-1519 (2000)
20211685
JOURNAL Direct Submission
Submitted (08-JUL-1999) Bioscience Division, Los Alamos National Laboratory, MS:M888, Los Alamos, NM 87545, USA
FEATURES
Source
Location/Qualifiers
1..975
/organism="Yersinia pestis"
/strain="Pestoides F"
/db_xref="taxon:632"
1..975
/gene="lcrv"
1..975
/gene="lcrv"
/codon_start=1
/transl_table=11
/product="V antigen"
/protein_id="AAF64076.1"
/db_xref="GI:7578512"
/translation="MIRAYEQNPQHIEDLEKRVQELTGHSGSYLEELVOLVYDKNT
DISIKYDPRKDESEFANVITDDIELLKRILAFPEPAALKGHYDQNLNGIKRV
EFLSESPYQWELRALFVAMHSLADIDDLKLVYDSDNMHGDASKELELAE
TAEIKYIVSIOALINKHLSGSGTINIHDKSNIMDKNLYGYTDETFPASAKEYLKE
MPTQIOWDSEKRTVSKIDPLGSNKTGALNKNKSYSTKNDNEISHRAITCSDK
SRPLNDLSQVTTQSDITSIFNSKATELNRFIOYDVSVMORLDVTR"
BASE COUNT 334 a 174 c 194 g 273 t
ORIGIN
Query Match 58.6%; Score 918; DB 2; Length 975;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 968; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 586 ATGATTAGAGCTTACGACAAACCAACCAACATTTATGAGAGCTAGAAAAAGTTAGG 645
DB 1 ATGATTAGAGCTTACGACAAACCAACCAACATTTATGAGAGCTAGAAAAAGTTAGG 60
QY 646 GTGGAACAACCTTACGTCATGCTCTTCAGTTTGAAGAAGTTGGTTCAGTTAGTCAAA 705
DB 61 GTGGAACAACCTTACGTCATGCTCTTCAGTTTGAAGAAGTTGGTTCAGTTAGTCAAA 120
QY 706 GATTAATAATGATATTTCCATTAATATGATCCCAAGAAAAGTTGCGAGTTTGGC 765
DB 121 GATTAATAATGATATTTCCATTAATATGATCCCAAGAAAAGTTGCGAGTTTGGC 180
QY 766 AATAGAGTAATTAATGATATGATATGATGCTCAAGAAAATCCTAGCTATTTTTCACCC 825
DB 181 AATAGAGTAATTAATGATATGATATGATGCTCAAGAAAATCCTAGCTATTTTTCACCC 240
QY 826 GAGATACCATTTCTTAAGGCGGTGATTAATGACAAACCAACTGCAAAATGCGATCAAGCA 885
DB 241 GAGATACCATTTCTTAAGGCGGTGATTAATGACAAACCAACTGCAAAATGCGATCAAGCA 300
QY 886 GTAAAGAGTTCCTTGATTCATCGCGGATATACCAATGGGAATTCGGGCGTTCATGGCA 945

```

```

DB 301 GTAAAGAGTTCCTTGATTCATCGCGGATATACCAATGGGAATTCGGGCGTTCATGGCA 360
QY 946 GATATGATTTCTCTTTAAACCGCGATGATATGATGATATTTGAAAGTATGTT 1005
DB 361 GATATGATTTCTCTTTAAACCGCGATGATATGATGATATTTGAAAGTATGTT 420
QY 1006 GATTCATGATATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
DB 421 GATTCATGATATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1066 ACCGCGAATTAAGATTTATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
DB 481 ACCGCGAATTAAGATTTATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 1126 ACTGCGACCATTAATATTCATGATATTAATCATGATGATGATGATGATGATGATGATGAT 1185
DB 541 ACTGCGACCATTAATATTCATGATATTAATCATGATGATGATGATGATGATGATGATGAT 600
QY 1186 TATACAGATGAAGATTTTAAAGCCAGCGAGTCAAAATTTCTGAGAAAATGCT 1245
DB 601 TATACAGATGAAGATTTTAAAGCCAGCGAGTCAAAATTTCTGAGAAAATGCT 660
QY 1246 CAACCCACCATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
DB 661 CAACCCACCATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 1306 GGAATGAGATTAAGAAAGAACCGGCGGCTGGTATCTGAATAAATCTCATCTATAT 1365
DB 721 GGAATGAGATTAAGAAAGAACCGGCGGCTGGTATCTGAATAAATCTCATCTATAT 780
QY 1366 AAAGATTAATTAATGATATTCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1425
DB 781 AAAGATTAATTAATGATATTCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 1426 AACGACTGTGTTAGCCAAACCAACTCAGCTGTGATATTAATCAATCACTTTATATCA 1485
DB 841 AACGACTGTGTTAGCCAAACCAACTCAGCTGTGATATTAATCAATCACTTTATATCA 900
QY 1486 GGTATTGAAGCACTAAGCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGTGTA 1545
DB 901 GGTATTGAAGCACTAAGCGTTTCATTCAGAAATATGATTCAGTATGCAACGCTGTGTA 960
QY 1546 GATGACACG 1554
DB 961 GATGACACG 969

```

RESULT 12

AF167310 981 bp DNA BCT 17-APR-2000

LOCUS AF167310

DEFINITION Yersinia pestis strain Angola V antigen (lcrv) gene, complete cds.

ACCESSION AF167310

VERSION AF167310.1 GI:7578514

KEYWORDS Yersinia pestis.

SOURCE Yersinia pestis.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.

REFERENCE 1 (bases 1 to 981)

AUTHORS Adair, D.M., Worsham, P.L., Hill, K.K., Klevytska, A.M., Jackson, P.J., Friedlander, A.M. and Keim, P.

Diversity in a variable-number tandem repeat from Yersinia pestis J. Clin. Microbiol. 38 (4), 1516-1519 (2000)

20211685

JOURNAL MEDLINE

REFERENCE 2 (bases 1 to 981)

AUTHORS Hill, K.K. and Jackson, P.J.

Diversity in a variable-number tandem repeat from Yersinia pestis J. Clin. Microbiol. 38 (4), 1516-1519 (2000)

20211685

JOURNAL Direct Submission

Submitted (08-JUL-1999) Bioscience Division, Los Alamos National Laboratory, MS:M888, Los Alamos, NM 87545, USA

FEATURES

Source

Location/Qualifiers

1..981

/organism="Yersinia pestis"

/strain="Angola"

gene
1. .981
/db.xref="taxon:632"
/gene="icrv"
1. .981
/gene="icrv"
/codon_start=1
/transl_table=1
/product="V antigen"
/protein_id="AF64077.1"
/db.xref="GI:7578515"
/translation="MIRAEONPOHIEDLEENVRDOLTGSSVLELYOYKDKNI
DISIKYPRKDSYFANRVITDDIELLRKILATFLPEDALIKGHTNOJONGIKRK
EFLESSPNTQWELAEVAVHFSLTADRIDDDILKVLVDSNHHGADRSKLEELAE
TAEIKYISVQAEINKHLSSSGYINIHDKNLGYNDEIRKASKEYLELLEK
MPQTTIOVDSEKIVSINKDFGSENKRTGALGKNSYKNNKNSLSPHATSSDK
SRPLNDIVDSKTTQLSDITSFNSAIEALNFIKXDSVWGRLLDDISGR"

BASE COUNT 334 a 175 c 196 g 276 t
ORIGIN

Query Match 49.68; Score 777; DB 2; Length 981;
Best Local Similarity 99.68; Pred. No. 0;
Matches 977; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 586 ATGATTAGAGCTTACGACAAACCAACACATTATTAGAGATCTAGAAAAGTTAGG 645
DB 1 ATGATTAGAGCTTACGACAAACCAACACATTATTAGAGATCTAGAAAAGTTAGG 60

QY 646 GTGGAACAACCTACTGCTGATGCTTCTCAGTTTAGAAGAAATGCTCAGTTAGTCAA 705
DB 61 GTGGAACAACCTACTGCTGATGCTTCTCAGTTTAGAAGAAATGCTCAGTTAGTCAA 120

QY 706 GATAAAATATAGATTTCCATTAAATATGATCCCAAGAAAGATTCCGAGGTTTGGC 765
DB 121 GATAAAATATAGATTTCCATTAAATATGATCCCAAGAAAGATTCCGAGGTTTGGC 180

QY 766 AATAGATAATTAAGTATGATATCGAATGCTCAAGAAATCTAGCTTATTTTACCC 825
DB 181 AATAGATAATTAAGTATGATATGATGCTCAAGAAATCTAGCTTATTTTACCC 240

QY 826 GAGATACCATCTTAAAGGGGCTATATGACAAACCACTGCAAAATGCGATCAAGCA 885
DB 241 GAGATACCATCTTAAAGGGGCTATATGACAAACCACTGCAAAATGCGATCAAGCA 300

QY 886 GTAAAGAGTCTTGAATCATCGCCGAATACAAATGGAATGCGGGCGTTATGGA 945
DB 301 GTAAAGAGTCTTGAATCATCGCCGAATACAAATGGAATGCGGGCGTTATGGA 360

QY 946 GTAATGCAATTTCTTTAACCGCGATGATGATGATATTTGAAAGTATGTT 1005
DB 361 GTAATGCAATTTCTTTAACCGCGATGATGATGATATTTGAAAGTATGTT 420

QY 1006 GATTCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
DB 421 GATTCATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

QY 1066 ACCGCCAATTAAGATTTATTCAGTATTCAGCCGAATTAATAGATCTGCTAGT 1125
DB 481 ACCGCCAATTAAGATTTATTCAGTATTCAGCCGAATTAATAGATCTGCTAGT 540

QY 1126 AGTGGCACCATAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
DB 541 AGTGGCACCATAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

QY 1186 TATACAGTGAAGATTTTAAAGCCAGCGAGTACAAAATTCGAGAAAATGCGCT 1245
DB 601 TATACAGTGAAGATTTTAAAGCCAGCGAGTACAAAATTCGAGAAAATGCGCT 660

QY 1246 CAACACCATTCAGTGGATGGAGCGAGAAAATAGTGTGATAAGGACTTCTT 1305
DB 661 CAACACCATTCAGTGGATGGAGCGAGAAAATAGTGTGATAAGGACTTCTT 720

QY 1306 GGAAGTGAATAAAGAACGGGGCGTTGGGTAATCTGAAAAACTCATCTTATTAAT 1365

DB 721 GGAAGTGAATAAAGAACGGGGCGTTGGGTAATCTGAAAAACTCATCTTATTAAT 780

QY 1366 AAGATATAATGATTAATTCCTACTTTGCCACACCGCTGCGATTAAGTCCAGGCCGCTC 1425
DB 781 AAGATATAATGATTAATTCCTACTTTGCCACACCGCTGCGATTAAGTCCAGGCCGCTC 840

QY 1426 AACGACTTGTTAGCCAAAACCAACACAGCTGCTGATATTAATCACTGTTTAAATCA 1485
DB 841 AACGACTTGTTAGCCAAAACCAACACAGCTGCTGATATTAATCACTGTTTAAATCA 900

QY 1486 GCTATGAGACGCTGACGACCGCTTCTATTCAGAAATATGATTCAGTATGACGACGCTGCTA 1545
DB 901 GCTATGAGACGCTGACGACCGCTTCTATTCAGAAATATGATTCAGTATGACGACGCTGCTA 960

QY 1546 GATGACACGCTGCTGATTAATGA 1566
DB 961 GATGACACGCTGCTGATTAATGA 981

RESULT 13
A56812
LOCUS A56812 547 bp DNA
DEFINITION Sequence 20 from Patent WO9628551.
ACCESSION A56812
VERSION A56812.1 GI:3712825
KEYWORDS
SOURCE
ORGANISM
Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
REFERENCE
AUTHORS 1 (bases 1 to 547)
Tilball, R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C., Bennett and Alice, M.
TITLE VACCINES FOR PLAGUE
JOURNAL Patent: WO 9628551-A 20 19-SEP-1996;
COMMENT OTHER PUBLICATION ZA 9602036 960716
Other publication AU 4951196 961002.
FEATURES
source
1. .547
/organism="Yersinia pestis"
/db.xref="taxon:632"
29. .538
/note="unnamed protein product"
CDS
/codon_start=1
/transl_table=1
/protein_id="CA03422.1"
/translation="MKRISSVIALEGTATNADLTASTATATATVAPRTLLY
KEBAPTIMNGNIDIELLYGTLGSKYKGTSTSVNPFDAAGDPRYLFTSDGNN
HOFRTKVGDSDFDISPRVNSENLVLDVYVATGSQDFVSHSGSKGLAAGRT
DAVTVYSN"

BASE COUNT 165 a 120 c 115 g 147 t
ORIGIN

Query Match 32.88; Score 513; DB 5; Length 547;
Best Local Similarity 100.0%; Pred. No. 2e-247;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 ATAGAAAAAATAGTCCGTATTCGCAATGATTAATTTGAACTATTCACATGCTA 127
DB 27 ATAGAAAAAATAGTCCGTATTCGCAATGATTAATTTGAACTATTCACATGCTA 86

QY 128 ATGGGCGAGATTAACTGCAAGCACAACCTGCAAGGCACTCTGTTGAAACAGCCGCA 187
DB 87 ATGGGCGAGATTAACTGCAAGCACAACCTGCAAGGCACTCTGTTGAAACAGCCGCA 146

QY 188 TCACCTTTACATATATAGAGAGCGCTCCCAATTAATATGAGCAATGAGAAACATCGATA 247
DB 147 TCACCTTTACATATATAGAGAGCGCTCCCAATTAATATGAGCAATGAGAAACATCGATA 206

QY 248 CAGAACTACTTGTGGTACGCTTCTTGGCGGCTATATAACAGACAGACAT 307
 DB 207 CAGAACTACTTGTGGTACGCTTCTTGGCGGCTATATAACAGACAGACAT 286
 QY 308 CAGTACTACTTACAGATGCGGCGGCGGATGCCATGTAATTAATTTACTTCAGAGATG 367
 DB 267 CAGTACTACTTACAGATGCGGCGGCGGATGCCATGTAATTAATTTACTTCAGAGATG 326
 QY 368 GAAATACACCAACCAATTCACCTACCAAAAGTATGGAAGATTTAGATATCT 427
 DB 327 GAAATACACCAACCAATTCACCTACCAAAAGTATGGAAGATTTAGATATCT 386
 QY 428 CTCCTAAGGTAACCGGTGAGACCTTGTGGGATGACGCTGCTTGGCTACGGCAGCC 487
 DB 387 CTCCTAAGGTAACCGGTGAGACCTTGTGGGATGACGCTGCTTGGCTACGGCAGCC 446
 QY 488 AGGATTTCTTGTGCTGCTCAATGTTGCAAAAGCGGTAACTGACAGCAAGTAATACA 547
 DB 447 AGGATTTCTTGTGCTGCTCAATGTTGCAAAAGCGGTAACTGACAGCAAGTAATACA 506
 QY 548 CTGATGCTGTAACCGGTACCGGTATCTAACCAAG 580
 DB 507 CTGATGCTGTAACCGGTACCGGTATCTAACCAAG 539

RESULT 14
 LOCUS YPCAF 5383 bp DNA BCT 23-NOV-1995
 DEFINITION Y.pestis genes cafI, cafIM, cafII and cafIR.
 ACCESSION K61996.540525.890405.X57773
 VERSION X61996.1.GI:48620
 KEYWORDS cafI gene; cafII gene; cafIM gene; cafIR gene; F1 antigen.
 SOURCE Yersinia pestis.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
 REFERENCE 1 (bases 1 to 5383)
 AUTHORS Gal'ov,E.E., Karlishhev,A.V., Chernovskaya,T.V., Dolgikh,D.A., Smirnov,O.Yu., Volkovoy,K.I., Abramov,V.M. and Zav'yalov,V.P.
 TITLE Expression of the envelope antigen F1 of Yersinia pestis is mediated by the product of cafIM gene having homology with the chaperone protein Papp of Escherichia coli
 JOURNAL FEBS Lett. 286 (1-2), 79-82 (1991)
 MEDLINE 91323540
 REFERENCE 2 (bases 1 to 5383)
 AUTHORS Gal'ov,E.E., Smirnov,O.Yu., Karlishhev,A.V., Volkovoy,K.I., Denesnyuk,A.I., Nazimov,I.V., Rubtsov,K.S., Abramov,V.M., Dalaydzanz,S.M. and Zav'yalov,V.P.
 TITLE Nucleotide sequence of the Yersinia pestis gene encoding F1 antigen and the primary structure of the protein. Putative T and B cell epitopes
 JOURNAL FEBS Lett. 277 (1-2), 230-232 (1990)
 MEDLINE 91099503
 REFERENCE 3 (bases 1 to 5383)
 AUTHORS Karlishhev,A.V.
 TITLE Direct Submission
 JOURNAL Submitted (24-JAN-1992) A.V. Karlishhev, Inst of Immunology, State Concern 'Biopreparation', 142380 Lyubuchany, Moscow Region, USSR revised by [51]
 REMARK 4 (bases 1 to 5383)
 REFERENCE Karlishhev,A.V., Gal'ov,E.E., Abramov,V.M. and Zav'yalov,V.P.
 AUTHORS CafIR gene and its role in the regulation of capsule formation of Y. pestis
 JOURNAL FEBS Lett. 305 (1), 37-40 (1992)
 MEDLINE 92339520
 REFERENCE 5 (bases 1 to 5383)
 AUTHORS Karlishhev,A.V.
 TITLE Direct Submission
 JOURNAL Submitted (22-SEP-1992) to the EMBL/GenBank/DBJ databases
 REFERENCE 6 (bases 1 to 5383)
 AUTHORS Karlishhev,A.V., Gal'ov,E.E., Smirnov,O.Yu., Guzeyev,A.P., Abramov,V.M. and Zav'yalov,V.P.
 TITLE A new gene of the F1 operon of Y. pestis involved in the capsule

JOURNAL biogenesis
 MEDLINE FEBS Lett. 297 (1-2), 77-80 (1992)
 FEATURES 92201398
 SOURCE Location/Qualifiers
 1..5383
 /organism="Yersinia pestis"
 /db_xref="taxon:632"
 complement(1..906)
 /gene="cafI"
 complement(1..906)
 /gene="cafIR"
 /function="positive regulator of f1 operon expression"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA43969.1"
 /db_xref="GI:48623"
 /db_xref="SWISS-PROT:P26950"
 /translation="MLKMTVNSIIIOYIEENLESEKFINIDCLVYSGFSRRYLOISFK EYVGMPICTYLRVRASRAALRLTLITIEISAKFLPYDSQOTFREFKRIFGYTPR OYRMIPFWSFKGLGRREINCEYLOPRICYLKERNIIGOCNFRDLVYSGIDSCLR GLYDSLKNTKNTAIVSNRIPEHDKNDIARTVWDRKHSDEIKVDGLVAYFF NDTYDQVHHMNYINSLPIYNLNKRBDGYVEVIRKNDNTIDCHYFLPYCDMER YNMGVYHNNTYKPEMSYTLGLPKS"
 1235..2011
 /gene="cafIM"
 1235..2011
 /gene="cafII"
 /function="chaperone-like protein involved in transport of F1 subunit"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA43967.1"
 /db_xref="GI:1072424"
 /db_xref="SWISS-PROT:P26926"
 /translation="MILNRLSTIGITITGMSPAAANSQPIPKRSKRYGTIGESRI ITPDASGVNSVKNTOYEPVLIOSRIIDENKKESEDEPVVTPPLEFLDKAQNLSR IAQAGVPRDRKESLKLWCIVGIPKPKDIDATNKNKPNPDVGEVFOAGVFINNC IKLLVRPELKGTPIOFAEKLKSWKVDGKLLAENSPYNNIGELTFGKSIPISHYIP PKSTAFPLPGLAGARVSWRIINDGGLRLYSKNVTL"
 2036..4537
 /gene="cafIA"
 2036..4537
 /gene="cafIA"
 /function="involved in anchoring capsule to cell surface"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA43968.1"
 /db_xref="GI:48622"
 /db_xref="SWISS-PROT:P26949"
 /translation="MRYSKFLCAGLITLPCMGRAATPDSMTLDTNSGSDISLF NQGLDPENYFVNVNGRKVDGNDIPRLKHNGKELMLPCLSLQUTKYGIDKY PDLTSGTECCVDLAIPIHSDVOYFENQOKSLIYVPPALLPRDGLIMDMDGIP ALFMYNNTNMOTRFRREGKSLDSVYNALOGCINGNARFSSISMKQCKOSYIT AERGINTKSRITLGETYSDSITPDSITPKIKIKIASDSKSPITOWNFAVVRGIART QARVEVLKDGITVSNELVPSGFELANPLDGGSGELVLIHESDGRQVTVYDYL AVALEKGYFESXSMGGEVPPANDLTOYSYVFGAKKYLPRNFTLYGLOSQSYHAE ALGIGAMGDFGALSTVDYQADSOKNOKKSGGOWRVRYNKYLOSGLINIAEEVA TEGEKKLADLNTYCKPNTNRDCRPDPAKPKNKYOENLSOSITGSGTINFGSKRWL RDSRSTFSYGVYHNEFRNGMSLTINLSTKNTINKRTSGKTESLISNITLSPLSRWLNNSINSNTOMTSDSHGNTTHEVGYGAEFDRQLTWDERENEGRKRTSVALNLNR GTYGEISGNYSYTDYQSOGLGVNGVNVYTOYGTAGKTDITLVAOPDISGASVG YWPKMDTFRERTYNTGYLTPYREKVENIPTLNDAEITNNIVSIVPTKAVYLAKE NARIIGRLFLHKRSDNKVPFGSIVTIEGSSSGIYGDNSGYVLTGLPKKSKILVK MGRDKNOSCSNVLPKRTDLSGARVRLSTIOILNN"
 4618..5130
 /gene="cafI"
 4618..5130
 /gene="cafI"
 /codon_start=1
 /transl_table=11
 /product="F1 capsule antigen"
 /protein_id="CAA43966.1"
 /db_xref="GI:48621"

Wed Aug 23 11:45:50 2000

us-08-699-716a-1.oligo.rge

Page 25

Search completed: August 22, 2000, 16:32:27
Job time: 4840 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 15:39:08 ; Search time 63.88 Seconds
(without alignments)
6133.388 Million cell updates/sec

Title: US-08-699-716a-1

Perfect score: 1566
Sequence: 1 ATGGGCATCATCATCATCA.....ATGACACGTCTGTGAATGA 1566

Scoring table: OLIGO_NMC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	927	59.2	1014	1 T04222	Partial LcrV (V an
2	923	58.9	1014	1 T38243	Y. pestis lcrV (V
3	876	55.9	1014	1 T38242	Y. pestis lcrV (V
4	876	55.9	1330	1 T38256	Y. pestis Fl/V ant
5	876	55.9	1530	1 T38249	Y. pestis Fl/V ant
6	872	55.7	1014	1 T04223	Partial LcrV (V an
7	513	32.8	547	1 T38248	Y. pestis Fl antiq
8	512	32.7	542	1 Q92819	Yersinia pestis ca
9	512	32.7	544	1 V41594	Nucleotide sequenc
10	512	32.7	544	1 V41594	Nucleotide sequenc
11	510	32.6	510	1 V41595	Nucleotide sequenc
12	448	28.6	450	1 V41601	Nucleotide sequenc
13	448	28.6	474	1 V41600	Nucleotide sequenc
14	447	28.5	447	1 V41609	Nucleotide sequenc
15	402	25.7	576	1 V41598	Nucleotide sequenc
16	398	25.4	541	1 Q92817	Yersinia pestis ca
17	398	25.4	541	1 T38244	Y. pestis catI (Fl
18	396	25.3	542	1 Q92818	Yersinia pestis ca
19	339	21.6	513	1 V41599	Nucleotide sequenc
20	68	4.3	375	1 V41599	DNA fragment 2 of
21	68	4.3	375	1 V41599	DNA fragment 3 of
22	68	4.3	375	1 V41599	DNA fragment 4 of
23	68	4.3	375	1 V41599	Chlamydomonas rein
24	53	3.4	1278	1 T91246	Type A neurotoxin
25	53	3.4	1402	1 V30572	Clostridium botuli
26	53	3.4	1460	1 V30593	Clostridium botuli
27	53	3.4	1460	1 V30584	Clostridium botuli
28	53	3.4	1469	1 V30591	Clostridium botuli
29	53	3.4	1472	1 V30585	Clostridium botuli
30	53	3.4	1502	1 V30588	Clostridium botuli
31	53	3.4	1526	1 V30580	Clostridium botuli
32	53	3.4	1535	1 V30596	Clostridium botuli
33	53	3.4	1546	1 V30575	Clostridium botuli

34	53	3.4	1547	1 V30581	Clostridium botuli
35	35	2.2	47	1 T37132	Sense primer T20-I
36	35	2.2	861	1 O81500	SVF anti-rev seque
37	35	2.2	861	1 T45347	Single chain SVF a
38	30	1.9	416	1 O59427	Human brain expres
39	30	1.9	8920	1 O62924	Cardiomyol-phospat
40	29	1.9	257	1 V88713	EST clone RH378. N
41	29	1.9	598	1 N90919	Type A insertion g
42	29	1.9	1021	1 T04613	5' flanking region
43	29	1.9	2359	1 Q58609	Sequence encoding
44	29	1.9	2787	1 X15661	Protein phosphatas
45	29	1.9	4459	1 V99094	DNA methyltransfer
46	29	1.9	4641	1 N81538	Sequence of a gene
47	29	1.9	19440	1 V99129	DNA methyltransfer
48	28	1.8	39	1 T05194	Human/mouse quanyl
49	28	1.8	39	1 X15374	Human quanylate ki
50	28	1.8	76	1 V08564	Primer for Transge
51	28	1.8	86	1 T17965	Plasmid pET15b His
52	28	1.8	287	1 T16383	pET-15b expression
53	28	1.8	552	1 V13947	Salmonella sefa ge
54	28	1.8	1247	1 X29725	Insert from pET28C
55	28	1.8	1310	1 V08558	Transgene for epit
56	28	1.8	2136	1 V30788	Topoisomerase I ex
57	28	1.8	3263	1 V08560	Transgene for epit
58	28	1.8	4286	1 V08559	Transgene for epit
59	27	1.7	38	1 V41618	Nucleotide sequenc
60	27	1.7	250	1 T22691	Human gene signatu
61	25	1.6	25	1 Q92816	DNA primer for Yer
62	25	1.6	25	1 T38247	Y. pestis Fl antiq
63	25	1.6	904	1 V06104	Vital infection ge
64	24	1.5	230	1 T24068	Human gene signatu
65	24	1.5	447	1 Q92652	Recombinant mite a
66	24	1.5	1667	1 O48589	Alcohol oxidase ge
67	23	1.5	30	1 V41617	Nucleotide sequenc
68	23	1.5	32	1 Q92814	DNA primer for Yer
69	23	1.5	32	1 T38245	Y. pestis Fl antiq
70	23	1.5	47	1 V58140	Cmb antisense PCR
71	23	1.5	53	1 V44450	Mycobacterium anti
72	23	1.5	53	1 V44407	Mycobacterium anti
73	23	1.5	53	1 V64559	M. tuberculosis fu
74	23	1.5	53	1 V64516	M. tuberculosis fu
75	23	1.5	135	1 T25107	Human gene signatu
76	23	1.5	553	1 X51720	DNA encoding a hum
77	23	1.5	797	1 V36459	cayAEI DNA sequenc
78	23	1.5	5434	1 V64410	Mouse llyor-1(gen
79	23	1.5	7635	1 Q27811	Sequence encoding
80	23	1.5	7635	1 O84653	Human neuronal cal
81	23	1.5	7635	1 O84654	DNA encoding human
82	23	1.5	7635	1 V42697	DNA encoding human
83	23	1.5	7635	1 V42697	DNA encoding human
84	23	1.5	7676	1 V55801	Mycobacterium anti
85	23	1.5	7676	1 V64567	M. tuberculosis fu
86	22	1.4	33	1 V41608	Nucleotide sequenc
87	22	1.4	45	1 V58142	Amp antisense PCR
88	22	1.4	1020	1 N92873	Expression plasmid
89	22	1.4	1518	1 N92872	Expression plasmid
90	22	1.4	1665	1 N92871	Expression plasmid
91	22	1.4	2079	1 N94527	Peptidyl C-termina
92	22	1.4	2625	1 N94527	Peptidyl C-termina
93	22	1.4	3315	1 N90791	Nucleotide sequenc
94	21	1.3	30	1 V41610	Nucleotide sequenc
95	21	1.3	37	1 V41610	Nucleotide sequenc
96	21	1.3	47	1 T72159	Reverse primer for
97	21	1.3	97	1 T29649	Primer for amplify
98	21	1.3	104	1 T33359	Ich-2 gene 5' prim
99	21	1.3	262	1 V89773	EST clone CT220. N
100	21	1.3	1477	1 O84557	Murine C140 recept
101	21	1.3	1477	1 T32036	Murine C140 recept
102	21	1.3	1497	1 T47066	Pinto bean alpha-D
103	21	1.3	1648	1 Q25532	Sequence of genom
104	21	1.3	2732	1 O84559	Murine C140 recept
105	21	1.3	2732	1 T32038	Murine C140 recept
106	20	1.3	20	1 O55602	5' flanking sequen

Clostridium botuli
Sense primer T20-I
SVF anti-rev seque
Single chain SVF a
Human brain expres
Cardiomyol-phospat
EST clone RH378. N
Type A insertion g
5' flanking region
Sequence encoding
Protein phosphatas
DNA methyltransfer
Sequence of a gene
DNA methyltransfer
Human/mouse quanyl
Human quanylate ki
Primer for Transge
Plasmid pET15b His
pET-15b expression
Salmonella sefa ge
Insert from pET28C
Transgene for epit
Topoisomerase I ex
Transgene for epit
Transgene for epit
Nucleotide sequenc
Human gene signatu
DNA primer for Yer
Y. pestis Fl antiq
Vital infection ge
Human gene signatu
Recombinant mite a
Alcohol oxidase ge
Nucleotide sequenc
DNA primer for Yer
Y. pestis Fl antiq
Cmb antisense PCR
Mycobacterium anti
Mycobacterium anti
M. tuberculosis fu
M. tuberculosis fu
Human gene signatu
DNA encoding a hum
cayAEI DNA sequenc
Mouse llyor-1(gen
Sequence encoding
O84653
Human neuronal cal
DNA encoding human
DNA encoding human
Mycobacterium anti
M. tuberculosis fu
Nucleotide sequenc
Amp antisense PCR
Expression plasmid
Expression plasmid
Expression plasmid
Peptidyl C-termina
Peptidyl C-termina
Nucleotide sequenc
Nucleotide sequenc
Reverse primer for
Primer for amplify
Ich-2 gene 5' prim
EST clone CT220. N
Murine C140 recept
Murine C140 recept
Pinto bean alpha-D
Sequence of genom
Murine C140 recept
Murine C140 recept
5' flanking sequen

C 107	20	1.3	28	1	T38258	Y. pestis F1 antiG	C 180	18	1.1	58	1	Q65621	C-terminal primer
C 108	20	1.3	40	1	Q55601	Flanking sequences	C 181	18	1.1	58	1	V64812	Zona pellucida zp
C 109	20	1.3	1133	1	X03474	Mollicella succinoy	C 182	18	1.1	63	1	Q47755	Globoblastoma deriv
C 110	20	1.3	1297	1	N81178	Plasmid pHCIFram1	C 183	18	1.1	195	1	T90307	Sequence used in d
C 111	20	1.3	1440	1	V47648	Fibroblast growth	C 184	18	1.1	537	1	T29650	ABC protease p II
C 112	20	1.3	2469	1	Q13311	Basic RGF receptor	C 185	18	1.1	688	1	Q67559	Retinoblastoma pro
C 113	20	1.3	2652	1	Q21003	flg receptor prote	C 186	18	1.1	802	1	Q43054	Cyn di clone 22 (C
C 114	20	1.3	2733	1	T31051	Human fibroblast g	C 187	18	1.1	1293	1	T94539	Tobacco cinnamoyl
C 115	20	1.3	2856	1	Q10448	Human basic fibrob	C 188	18	1.1	1821	1	Q75408	Ovalbumin pl+H6ov
C 116	20	1.3	3328	1	Q20914	Human bFGF recepto	C 189	18	1.1	1881	1	Q75410	HCV NS3 DNA. New h
C 117	20	1.3	3525	1	Q27658	N-sam cDNA. Gene p	C 190	18	1.1	1933	1	X23258	Urochloa panicleide
C 118	20	1.3	5194	1	X25885	C.albicans alpha-1	C 191	18	1.1	2107	1	T11024	DNA encoding recom
C 119	19	1.2	33	1	V56369	Human ICAM-R cDNA	C 192	18	1.1	2142	1	T73930	Urochloa panicleide
C 120	19	1.2	40	1	V56370	Human ICAM-R cDNA	C 193	18	1.1	2160	1	T73931	DNA for recombinan
C 121	19	1.2	40	1	X27597	Primer for amplify	C 194	18	1.1	2220	1	T11022	Urochloa panicleide
C 122	19	1.2	41	1	Q93962	Japanese oyster tr	C 195	18	1.1	2722	1	Q94644	Recombinant cold-r
C 123	19	1.2	41	1	V40245	Human VEGF-B166 p	C 196	18	1.1	2722	1	T13675	E. histolytica pyr
C 124	19	1.2	49	1	Q71813	Interleukin-18 bin	C 197	18	1.1	2809	1	Q91287	ORC2 gene encoding
C 125	19	1.2	50	1	X24790	PSORT-id fragment	C 198	18	1.1	2809	1	T73280	S. cerevisiae orig
C 126	19	1.2	50	1	X24790	Interleukin-18 bin	C 199	18	1.1	3835	1	Q80004	Sucrose-phosphate-
C 127	19	1.2	53	1	T80835	Gene specific zip	C 200	18	1.1	3658	1	Q36874	Fragment of TROVAC
C 128	19	1.2	53	1	V47595	Rhodococcus haloi	C 201	18	1.1	3659	1	Q67858	Foxpox PuuII-EcoR
C 129	19	1.2	179	1	X19099	Oropouche NP cDNA	C 202	18	1.1	3659	1	T01469	Plasmid pRW31.15
C 130	19	1.2	179	1	V86242	EST clone AA287. N	C 203	18	1.1	3659	1	T00518	TROVAC DNA fragmen
C 131	19	1.2	466	1	T46174	Modified hsp60 gen	C 204	18	1.1	3659	1	T04699	PRW31.15 (contg.F
C 132	19	1.2	740	1	V37345	Streptococcus pneu	C 205	18	1.1	3660	1	Q45687	Recombinant fowlp
C 133	19	1.2	762	1	V04627	Flea cysteine prot	C 206	18	1.1	3660	1	V26865	TROVAC PuuII-EcoR
C 134	19	1.2	783	1	V04628	Flea cysteine prot	C 207	18	1.1	3660	1	V23255	Fowlpox PuuII-EcoR
C 135	19	1.2	811	1	V04625	DNA encoding His-c	C 208	18	1.1	3661	1	Q28108	Fragment of TROVAC
C 136	19	1.2	819	1	V83967	Flea cysteine prot	C 209	18	1.1	3661	1	Q36873	Fowlpox PuuII frag
C 137	19	1.2	823	1	V04626	HSV glycoprotein D	C 210	18	1.1	4453	1	V54118	Human cyclin D1-cy
C 138	19	1.2	927	1	T51517	Rhodococcus haloi	C 211	18	1.1	4453	1	T69890	Human cyclin D1-cy
C 139	19	1.2	973	1	V26250	Fibroblast growth	C 212	18	1.1	4540	1	T69891	Human cyclin D1-cy
C 140	19	1.2	1056	1	V09256	Nucleotide sequenc	C 213	18	1.1	4540	1	V54119	Human cyclin D1-cy
C 141	19	1.2	1314	1	V09256	Recombinant papill	C 214	18	1.1	6327	1	T32301	Dermatophytosis sp
C 142	19	1.2	1599	1	T05634	Arabidopsis thalia	C 215	18	1.1	6387	1	Q36498	Chimeric cytochrom
C 143	19	1.2	1636	1	T91940	Arabidopsis thalia	C 216	18	1.1	6387	1	T30354	Plasmid pET-BS(+)
C 144	19	1.2	1642	1	T91939	Arabidopsis thalia	C 217	18	1.1	7559	1	X23260	Plasmid pET-BS(+)
C 145	19	1.2	1642	1	T91941	Arabidopsis thalia	C 218	18	1.1	8145	1	X23259	Plasmid pET-BS(+)
C 146	19	1.2	1643	1	T91938	Human cytomagelov	C 219	18	1.1	19250	1	V52168	Streptococcus pneu
C 147	19	1.2	1896	1	T79194	Arabidopsis thalia	C 220	18	1.1	19390	1	V52219	Streptococcus pneu
C 148	19	1.2	1964	1	T91937	Modified doxa gene	C 221	18	1.1	3585	1	X20251	Borrelia burgdorfe
C 149	19	1.2	2081	1	V01452	B. pallidus DNA po	C 222	18	1.1	27	1	Q40996	Annealing oligonuc
C 150	19	1.2	3251	1	V11724	Human secretory pr	C 223	18	1.1	29	1	T04224	Primer for LcrV (V
C 151	19	1.2	3704	1	V54590	B. pallidus DNA po	C 224	18	1.1	41	1	T05193	Human/mouse quanyl
C 152	19	1.2	3831	1	T59450	Human papillomavir	C 225	18	1.1	124	1	X15373	Human gene signatu
C 153	19	1.2	4770	1	T59450	Human RFP-beta. H	C 226	18	1.1	247	1	V90220	Human gene signatu
C 154	19	1.2	6924	1	O67283	EP-892047 Seq ID 3	C 227	18	1.1	281	1	V89456	EST clone CO36. N
C 155	19	1.2	7108	1	X02662	Cytotactin gene. C	C 228	18	1.1	303	1	Q59440	Human brain expres
C 156	19	1.2	7286	1	T14547	EP-892047 Seq ID 3	C 229	18	1.1	294	1	T33716	Cellulobrevin-3 codi
C 157	19	1.2	7475	1	X02659	EP-892047 Seq ID 3	C 230	18	1.1	303	1	T33716	Human gene signatu
C 158	19	1.2	8192	1	X02660	Chimeric protease	C 231	18	1.1	336	1	T20814	Human brain expres
C 159	19	1.2	11326	1	T12373	Purification tag o	C 232	18	1.1	437	1	Q55395	Human V3 loop HIV
C 160	18	1.1	18	1	T71737	CDNA encoding a pe	C 233	18	1.1	933	1	V71744	Mouse pheromone re
C 161	18	1.1	18	1	V99571	Y. pestis V antige	C 234	18	1.1	1042	1	V18442	Unknown DNA sequen
C 162	18	1.1	28	1	T38250	DNA primer for ver	C 235	18	1.1	1068	1	N71176	Sequence encoding
C 163	18	1.1	30	1	Q92824	Primer DN517 ampl	C 236	18	1.1	1138	1	V46314	Human secreted pro
C 164	18	1.1	30	1	T66435	Zipper adapter U-F	C 237	18	1.1	1138	1	Q70736	Human TARA-binding
C 165	18	1.1	35	1	Q78544	Gene specific zip	C 238	18	1.1	1138	1	V46314	TARA-binding prote
C 166	18	1.1	35	1	T80833	Gene specific zip	C 239	18	1.1	1138	1	Q70736	Mouse pheromone re
C 167	18	1.1	40	1	V67161	S. avidin1 strept	C 240	18	1.1	1183	1	T42222	Potato tuber PPO c
C 168	18	1.1	43	1	X23354	H. virescens JHE P	C 241	18	1.1	1183	1	T79592	UNC Clone 6 (Amb a
C 169	18	1.1	45	1	X23348	3' primer for Deng	C 242	18	1.1	1236	1	X05852	Ragweed pollen UNC
C 170	18	1.1	45	1	T65108	3' primer for Deng	C 243	18	1.1	1236	1	X05849	Ragweed allergen U
C 171	18	1.1	48	1	T65107	PCR primer for Deng	C 244	18	1.1	1308	1	X05848	Sequence encoding
C 172	18	1.1	49	1	Q50503	PCR primer DKG-MOT	C 245	18	1.1	1325	1	Q36666	Full length Amb a
C 173	18	1.1	49	1	V19526	Wild type Salmonel	C 246	18	1.1	1328	1	V13356	Ragweed Pollen All
C 174	18	1.1	51	1	Q88136	Primer used in the	C 247	18	1.1	1338	1	N70758	
C 175	18	1.1	51	1	T50968	Methionase gene	C 248	18	1.1	1349	1	Q06223	
C 176	18	1.1	51	1	X23215	P. putida methion	C 249	18	1.1	1349	1	Q51240	
C 177	18	1.1	51	1	X28101	PCR primer for Met	C 250	18	1.1	1349	1		
C 178	18	1.1	51	1		Primer for amplifiy	C 251	18	1.1	1349	1		
C 179	18	1.1	53	1	T35905		C 252	18	1.1	1349	1	T33615	

C 253	17	1.1	1349	1	V13352	Ragweed pollen Amb	326	16	1.0	748	1	X24234	MO9916900 Seq ID 4
C 254	17	1.1	1349	1	V37861	Allergen Amb A ₁ IB	327	16	1.0	748	1	X24250	MO9916894 Seq ID 4
C 255	17	1.1	1434	1	X16007	DNA encoding an an	328	16	1.0	844	1	V52660	Human native hepat
C 256	17	1.1	1506	1	X20639	Polynucleotide seq	329	16	1.0	918	1	V24860	H. pylori secreted
C 257	17	1.1	1521	1	X05850	Mouse pheromone re	330	16	1.0	959	1	V04200	PT1913 used in dev
C 258	17	1.1	1635	1	X13558	Enterococcus faeca	331	16	1.0	966	1	V24619	H. pylori ORF 02ae
C 259	17	1.1	1670	1	X05132	Potato polyphenol	332	16	1.0	1016	1	070049	Dicrofilaria immiti
C 260	17	1.1	2188	1	Q98952	g19 insert encodin	333	16	1.0	1022	1	X30789	Streptococcus pneu
C 261	17	1.1	2226	1	V48180	Artichoke sucrose	334	16	1.0	1025	1	V18441	Unknown DNA sequen
C 262	17	1.1	2284	1	052700	Oestrogen-binding	335	16	1.0	1033	1	T58282	Arabidopsis STO po
C 263	17	1.1	2295	1	T31994	Nonsense-mediated	336	16	1.0	1042	1	Q06218	UNC Clone 15 (Amb
C 264	17	1.1	2409	1	X05842	Mouse pheromone re	337	16	1.0	1158	1	T49314	Hypersensitive res
C 265	17	1.1	2424	1	X05841	Mouse pheromone re	338	16	1.0	1166	1	T72146	Pea type I LhcIIb
C 266	17	1.1	2550	1	X05840	Mouse pheromone re	339	16	1.0	1184	1	Q94057	Strain p69p1 acar
C 267	17	1.1	2561	1	X05808	Mouse pheromone re	340	16	1.0	1185	1	Q20338	B. thuringiensis to
C 268	17	1.1	2732	1	X05810	Mouse pheromone re	341	16	1.0	1185	1	Q30807	Toxin 69D1. New Ba
C 269	17	1.1	2734	1	X05809	Mouse pheromone re	342	16	1.0	1185	1	Q30939	Encodes BT toxin 6
C 270	17	1.1	2773	1	X05813	Mouse pheromone re	343	16	1.0	1185	1	Q32163	Bt toxin 69D1 Nema
C 271	17	1.1	2821	1	X05812	Mouse pheromone re	344	16	1.0	1185	1	Q51691	Bacillus thuringie
C 272	17	1.1	2907	1	X05803	Mouse pheromone re	345	16	1.0	1185	1	X26320	Fusion of cry6B an
C 273	17	1.1	2956	1	V80739	Human secreted pro	346	16	1.0	1189	1	Q40203	Sequence of UFS192
C 274	17	1.1	2961	1	X05802	Mouse pheromone re	347	16	1.0	1189	1	V27350	Streptococcus pneu
C 275	17	1.1	2962	1	X05811	Mouse pheromone re	348	16	1.0	1196	1	N81594	Encodes Tobacco Ga
C 276	17	1.1	3074	1	Q66455	Human thymospond	349	16	1.0	1203	1	Q87971	Factor-XA cDNA. In
C 277	17	1.1	3080	1	X05801	Mouse pheromone re	350	16	1.0	1227	1	V63788	Recombinant gene e
C 278	17	1.1	3175	1	X24384	Arabidopsis farnes	351	16	1.0	1285	1	X13941	H. pylori GHP0 510
C 279	17	1.1	3707	1	Q82792	TRK1 gene. New tra	352	16	1.0	1287	1	Q55751	Erwinia amylovora
C 280	17	1.1	3983	1	N90569	Ovine Y-chromosoma	353	16	1.0	1288	1	V36428	Hypersensitive res
C 281	17	1.1	4080	1	T31993	Nonsense-mediated	354	16	1.0	1288	1	V39973	Erwinia amylovora
C 282	17	1.1	4464	1	N71244	Genomic sequence o	355	16	1.0	1288	1	V54607	DNA encoding a hyp
C 283	17	1.1	4465	1	Q28697	Gltamine syntheta	356	16	1.0	1288	1	V83989	PROLAC FLAG-connec
C 284	17	1.1	4616	1	V26971	Feline immunodef	357	16	1.0	1300	1	V33371	PROLAC FLAG-connec
C 285	17	1.1	5868	1	T18030	Human R12 allele D	358	16	1.0	1320	1	V33372	Amb a IC clone enc
C 286	17	1.1	8561	1	X06982	Candida albicans h	359	16	1.0	1320	1	Q06224	Ragweed Pollen A11
C 287	17	1.1	9280	1	V74442	Staphylococcus aur	360	16	1.0	1320	1	Q51241	Ragweed Amb ai. 3 a
C 288	17	1.1	9471	1	V57356	Feline immunodef	361	16	1.0	1320	1	T32616	Ragweed pollen Amb
C 289	17	1.1	9472	1	V25601	FIV-NC8U1 clone JS	362	16	1.0	1320	1	V13353	Allergen Amb a IC
C 290	17	1.1	9474	1	V58053	FIV immunodef	363	16	1.0	1320	1	T37862	peptidylglycine a1
C 291	17	1.1	9474	1	V57355	FIV F14 clone. Pro	364	16	1.0	1390	1	Q13751	Human cartilage 11
C 292	17	1.1	9474	1	V57355	FIV-NC8U1 clone JS	365	16	1.0	1400	1	Q70142	Staphylococcus aur
C 293	17	1.1	13875	1	Q20240	Human centromeric	366	16	1.0	1539	1	V74518	Plant SAR gene p66
C 294	17	1.1	17250	1	X13020	Enterococcus faeca	367	16	1.0	1697	1	Q99785	Tobacco SAR CHX 1n
C 295	17	1.1	36335	1	Q68003	Ad2-ORF6/PGK-CFTR	368	16	1.0	1697	1	V81684	Streptococcus pneu
C 296	17	1.1	58840	5	T58840_5	Continuation (6 of	369	16	1.0	1697	1	V62800	Tyrosine phenol-1y
C 297	17	1.1	110007	1	V21209_03	Continuation (4 of	370	16	1.0	1755	1	V43025	Enterococcus faeca
C 298	16	1.0	24	1	T37133	Antisense primer T	371	16	1.0	1801	1	Q04271	Fibrinogen binding
C 299	16	1.0	33	1	V41602	Nucleotide sequenc	372	16	1.0	1908	1	X12964	Mouse c-rafi-1 gene
C 300	16	1.0	36	1	V33370	Nucleotide sequenc	373	16	1.0	1938	1	T08079	Mouse c-rafi-1 gene
C 301	16	1.0	40	1	V33370	Proactin leader s	374	16	1.0	1947	1	Q23458	Human phosphodiesterase
C 302	16	1.0	42	1	X23255	H. vitescens JHE p	375	16	1.0	1947	1	T68801	Human phosphodiesterase
C 303	16	1.0	42	1	X23249	H. vitescens JHE p	376	16	1.0	1947	1	X03926	Human phosphodiesterase
C 304	16	1.0	45	1	V33391	Nucleotide sequenc	377	16	1.0	2063	1	V43601	Human phosphodiesterase
C 305	16	1.0	51	1	V53860	Plasmid vector pTR	378	16	1.0	2091	1	T90951	Human B-raf kinase
C 306	16	1.0	61	1	T34881	GFR expression pla	379	16	1.0	2091	1	T93221	Human foetal brain
C 307	16	1.0	76	1	T90806	Mouse persepilin ge	380	16	1.0	2119	1	X13605	Human PIGR-1 encod
C 308	16	1.0	107	1	V75755	Staphylococcus aur	381	16	1.0	2162	1	T34621	Enterococcus faeca
C 309	16	1.0	108	1	N60280	Sequence encoding	382	16	1.0	2171	1	V24426	Human phosphodiesterase
C 310	16	1.0	146	1	N60861	Fragment of plasmid	383	16	1.0	2229	1	Q23462	Human telomerase r
C 311	16	1.0	212	1	V75738	Staphylococcus aur	384	16	1.0	2343	1	V99727	Human B-raf protei
C 312	16	1.0	220	1	Q78811	HMV antibody heav	385	16	1.0	2345	1	X18951	Human foetal brain
C 313	16	1.0	224	1	T34867	Plasmid pBN2:GRF(1	386	16	1.0	2359	1	V07901	Cobra venom mocrath
C 314	16	1.0	366	1	T23514	Human gene signatu	387	16	1.0	2467	1	V17812	Mus musculus don-1
C 315	16	1.0	369	1	T34868	Human gene signatu	388	16	1.0	2510	1	X21931	Human B-raf kinase
C 316	16	1.0	398	1	T21837	Human gene signatu	389	16	1.0	2610	1	N93060	cDNA of plasmid px
C 317	16	1.0	412	1	V28851	Targeting vector T	390	16	1.0	2790	1	Q45353	Human protein-tyro
C 318	16	1.0	425	1	T35124	Enhanced expressio	391	16	1.0	3058	1	T77137	Single chain anti
C 319	16	1.0	447	1	Q92652	Recombinant mite a	392	16	1.0	3076	1	V43674	Receptor type tyro
C 320	16	1.0	480	1	V88441	TCF-beta-like cyto	393	16	1.0	3171	1	V32920	Solanum tuberosum
C 321	16	1.0	480	1	T48445	EST clone CV410. N	394	16	1.0	3193	1	X19252	Barley dihydroquer
C 322	16	1.0	513	1	T48450	TCF-beta-like cyto	395	16	1.0	3383	1	Q20269	Gene encoding Ae-I
C 323	16	1.0	526	1	V33517	Clone 297 partial	396	16	1.0	3393	1	V74737	Staphylococcus aur
C 324	16	1.0	604	1	T27033	Camphylobacter coli	397	16	1.0	3441	1	T87922	Rat cerebellum der
C 325	16	1.0	677	1	T97157	SV40 poly A signal	398	16	1.0	3704	1	T59938	Phage resistance g

c 399	16	1.0	4004	T51260	Human AD4 gene gen	c 472	15	1.0	384	1	T20972	Human gene signatu
c 400	16	1.0	4259	V04201	Merged contigs pfr	c 473	15	1.0	399	1	T44935	Partial sequence o
c 401	16	1.0	4317	V52257	Streptococcus pneu	c 474	15	1.0	419	1	T30731	Mouse cryptidin 6 g
c 402	16	1.0	4464	N71244	Genomic sequence o	c 475	15	1.0	422	1	T30740	Mouse cryptidin 2 c
c 403	16	1.0	4465	Q28697	Glutamine syntheta	c 476	15	1.0	422	1	T30744	Mouse cryptidin 6 c
c 404	16	1.0	5663	X12989	Enterococcus faeca	c 477	15	1.0	435	1	V70865	Internal transcrib
c 405	16	1.0	6008	Q41289	Ubiquitin-specific	c 478	15	1.0	426	1	V75379	Streptococcus pneu
c 406	16	1.0	6516	V52238	Streptococcus pneu	c 479	15	1.0	445	1	T19906	Human gene signatu
c 407	16	1.0	6619	V63789	Plasmid pKM72/68 e	c 480	15	1.0	468	1	X21097	Polynucleotide seq
c 408	16	1.0	6761	X20517	Polynucleotide seq	c 481	15	1.0	478	1	T83960	DNA encoding a Sta
c 409	16	1.0	7296	X12996	Enterococcus faeca	c 482	15	1.0	533	1	Q27947	CD44. Anther-spec
c 410	16	1.0	8133	T29248	C. difficile toxin	c 483	15	1.0	534	1	T64630	G-CSF receptor ago
c 411	16	1.0	8133	V30560	Clostridium diffic	c 484	15	1.0	540	1	Q13860	B.campestris seed-
c 412	16	1.0	10023	V52731	Human hepatocyte n	c 485	15	1.0	581	1	Q05760	CDNA clone pms14 c
c 413	16	1.0	10897	T09187	Mutu putative onco	c 486	15	1.0	581	1	Q08573	CDNA clone pms14 c
c 414	16	1.0	11191	X20578	Polynucleotide seq	c 487	15	1.0	581	1	Q08545	CDNA clone pms14 c
c 415	16	1.0	12011	V83540	PCR-generated regi	c 488	15	1.0	581	1	T60953	Tapetum-specific c
c 416	16	1.0	13542	V74465	Staphylococcus aur	c 489	15	1.0	581	1	Q05668	pms14 contig. male
c 417	16	1.0	14078	V74502	Staphylococcus aur	c 490	15	1.0	608	1	V17547	Peptidyl prollyl ci
c 418	16	1.0	14516	X06748	E. coli O111 anti	c 491	15	1.0	614	1	Q34547	IL-6 sss muten.
c 419	16	1.0	15462	V18272	HP1V-3 JS isolate	c 492	15	1.0	614	1	Q89361	Human interleukin-
c 420	16	1.0	15462	V18273	HP1V-3 FR1 cp45 v	c 493	15	1.0	614	1	T39895	Coding sequence fo
c 421	16	1.0	15462	V18274	HP1V-3 Vero cp45 v	c 494	15	1.0	615	1	T67398	H. pylori cytoplas
c 422	16	1.0	15462	V70401	Human parainfluenz	c 495	15	1.0	633	1	V72021	Adenovirus T51gDE
c 423	16	1.0	15462	V83561	Nucleotide sequenc	c 496	15	1.0	645	1	V74997	Staphylococcus aur
c 424	16	1.0	15660	V83532	Plasmid p3/7(131)	c 497	15	1.0	658	1	T36735	Enterococcus faeca
c 425	16	1.0	15666	V83533	Plasmid p3/7(131)2	c 498	15	1.0	686	1	X19869	DNA encoding a S.
c 426	16	1.0	15669	V83520	Plasmid p218(131)	c 499	15	1.0	686	1	X14334	H. pylori GHPO 212
c 427	16	1.0	19718	V52232	Streptococcus pneu	c 500	15	1.0	692	1	V74448	Staphylococcus aur
c 428	16	1.0	35524	V22140	Chimpanzee adenove	c 501	15	1.0	709	1	X30726	Streptococcus pneu
c 429	16	1.0	45564	X23520	Human kidney amino	c 502	15	1.0	717	1	T63298	Human proteasome a
c 430	16	1.0	110000	V21209_00	Methanococcus jamn	c 503	15	1.0	777	1	X14447	H. pylori GHPO 902
c 431	16	1.0	110000	V21209_10	Continuation (11 o	c 504	15	1.0	781	1	V06144	Viral infection ge
c 432	16	1.0	133894	T13635	ACNPV genomic DNA	c 505	15	1.0	786	1	T13721	ACNPV ORF 141, res
c 433	15	1.0	19	V35789	Human epimorphin c	c 506	15					

545	15	1.0	1369	1	V42918	Truncated Interleu	618	15	1.0	2372	1	V04836	CDNA sequence of h
546	15	1.0	1424	1	V74877	Staphylococcus aur	619	15	1.0	2372	1	V03607	CDNA sequence of h
547	15	1.0	1473	1	O87724	Human auxiliary cy	620	15	1.0	2372	1	V20549	Human MDM2 encodin
548	15	1.0	1473	1	O87725	Human auxiliary cy	621	15	1.0	2372	1	V20876	Human MDM2 gene. I
549	15	1.0	1473	1	O87726	Human auxiliary cy	622	15	1.0	2372	1	X03847	Human MDM2 encodin
550	15	1.0	1473	1	T17394	Human derived cyto	623	15	1.0	2372	1	X35093	Nucleotide sequenc
551	15	1.0	1473	1	T28390	Human cytochrome p	624	15	1.0	2372	1	T58328	Thermotoga maritima
552	15	1.0	1473	1	T28392	Human cytochrome p	625	15	1.0	2407	1	X22557	Human HPIA041 CDNA
553	15	1.0	1473	1	X23749	Rice anthranilate	626	15	1.0	2427	1	O11727	Sequence encoding
554	15	1.0	1498	1	O83682	Epsilon oploid rec	627	15	1.0	2427	1	O70435	Human glyco-protein
555	15	1.0	1518	1	O83682	Streptococcus pneu	628	15	1.0	2433	1	T84441	Staphylococcal MHC
556	15	1.0	1561	1	V42993	Human creatine kin	629	15	1.0	2440	1	O95542	NPW/ALK fusion gen
557	15	1.0	1568	1	O51620	Mutant human creat	630	15	1.0	2507	1	T98534	Natural splice var
558	15	1.0	1568	1	O58990	S10-95 CDNA clone	631	15	1.0	2527	1	O67556	Chitinase I gene.
559	15	1.0	1622	1	O68829	Staphylococcus aur	632	15	1.0	2558	1	V53514	DNA encoding a Sta
560	15	1.0	1623	1	V74810	S10-87 CDNA clone	633	15	1.0	2608	1	O93541	Partial ALK gene.
561	15	1.0	1642	1	O68828	DNA encoding epide	634	15	1.0	2624	1	T98532	Coding sequence fo
562	15	1.0	1643	1	O27436	Hepatitis E virus	635	15	1.0	2635	1	T95763	Arabidopsis SCAREC
563	15	1.0	1649	1	V61686	Human temporal lob	636	15	1.0	2877	1	T12913	Haematopoietin rec
564	15	1.0	1649	1	O14638	Plasmod pPDE2R in	637	15	1.0	2877	1	T95781	Human OB-R variant
565	15	1.0	1649	1	T34379	Staphylococcus aur	638	15	1.0	2880	1	T12912	Haematopoietin rec
566	15	1.0	1674	1	V74520	Adenovirus PAC1551	639	15	1.0	2880	1	T74022	Variant form of hu
567	15	1.0	1704	1	V72022	Schistosaccharomyce	640	15	1.0	2880	1	T95780	Human OB-R variant
568	15	1.0	1705	1	O79621	DNA encoding a X-P	641	15	1.0	2882	1	X20292	Borrelia burgdorfe
569	15	1.0	1709	1	T84145	DNA encoding one u	642	15	1.0	2907	1	V71026	p53alpha-green flo
570	15	1.0	1709	1	V53488	Secreted protein C	643	15	1.0	2913	1	V71035	p53alpha-green flo
571	15	1.0	1731	1	T84446	Rat STCH chaperone	644	15	1.0	2936	1	T65461	Thermophilic alkali
572	15	1.0	1762	1	X33193	Cytochrome P450 2C	645	15	1.0	2948	1	T98533	Coding sequence fo
573	15	1.0	1839	1	T11375	Human cytochrome p	646	15	1.0	2948	1	X18940	Human basic helix-
574	15	1.0	1839	1	V44156	CDNA of endothelia	647	15	1.0	2974	1	O40777	Genomic DNA sequen
575	15	1.0	1838	1	N90769	Staphylococcus aur	648	15	1.0	2976	1	T98530	Coding sequence fo
576	15	1.0	1847	1	V74678	Papaya ACC synthase	649	15	1.0	2991	1	T12911	Haematopoietin rec
577	15	1.0	1888	1	V31482	R. rubrum poly-bet	650	15	1.0	2991	1	T89193	Human OB-R lepton
578	15	1.0	1924	1	V80358	Interferon-omega-1	651	15	1.0	2991	1	T72649	Human OB-R lepton
579	15	1.0	1934	1	N60201	Nitrosomonas dhak	652	15	1.0	2991	1	T75172	Human haematopoiet
580	15	1.0	1935	1	V23811	Human secreted pro	653	15	1.0	3000	1	O92913	Human OB-R variant
581	15	1.0	1939	1	X04312	Hepatitis E virus	654	15	1.0	3000	1	X02913	Human protein tyro
582	15	1.0	1983	1	V61688	Hepatitis E virus	655	15	1.0	3000	1	V04696	Homo sapiens 20q13
583	15	1.0	1984	1	O86592	Hepatitis E virus	656	15	1.0	3004	1	T85576	Human MSX receptor
584	15	1.0	2030	1	T43203	Human mitogen-acti	657	15	1.0	3012	1	X05825	Rat pheromone rece
585	15	1.0	2030	1	X07065	Human HCNV inducib	658	15	1.0	3060	1	T62781	Urea transporter p
586	15	1.0	2056	1	X133945	B. campestris seed-	659	15	1.0	3063	1	T98529	Coding sequence fo
587	15	1.0	2060	1	O13861	Bce4 gene. New pla	660	15	1.0	3102	1	T85577	Human MSX receptor
588	15	1.0	2060	1	O22615	Bce4 genomic clone	661	15	1.0	3108	1	V20869	Human phospholipas
589	15	1.0	2060	1	O64002	Bce4 genomic clone	662	15	1.0	3108	1	X05814	Mouse pheromone re
590	15	1.0	2081	1	O13862	Neuronal alpha-bun	663	15	1.0	3125	1	X05805	Mouse pheromone re
591	15	1.0	2101	1	T59197	Chitinase 2 gene.	664	15	1.0	3125	1	X05805	Mouse pheromone re
592	15	1.0	2119	1	O56757	Human HPIA041 CDNA	665	15	1.0	3144	1	V52357	Streptococcus pneu
593	15	1.0	2156	1	X22558	Plasmod pPRAPDP in	666	15	1.0	3165	1	T84001	DNA encoding a Sta
594	15	1.0	2156	1	O14624	Plasmod pPRAPDP in	667	15	1.0	3165	1	V53408	DNA encoding 2 Sta
595	15	1.0	2158	1	T34372	Genomic sequence e	668	15	1.0	3167	1	T73392	Coding sequence fo
596	15	1.0	2159	1	O06958	Sequence of the pr	669	15	1.0	3167	1	T98528	Chlamydia pneumoni
597	15	1.0	2169	1	N30016	Streptococcus pneu	670	15	1.0	3222	1	V20867	Human phospholipas
598	15	1.0	2196	1	T04075	Tomato S-ribonucle	671	15	1.0	3222	1	X06816	Protein PRO332 CDN
599	15	1.0	2219	1	V37346	Human p21-activate	672	15	1.0	3222	1	X52367	Human phospholipas
600	15	1.0	2241	1	T64552	Human p21-activate	673	15	1.0	3296	1	O13967	Human p21-activate
601	15	1.0	2245	1	T11023	Human secreted pro	674	15	1.0	3307	1	X05822	Rat pheromone rece
602	15	1.0	2248	1	T29163	Human p21-activate	675	15	1.0	3359	1	X05824	Rat pheromone rece
603	15	1.0	2248	1	T47035	Human p21-activate	676	15	1.0	3372	1	O27354	Growth factor rece
604	15	1.0	2248	1	T99252	Human p21-activate	677	15	1.0	3425	1	O57072	Human GFR-1 genom
605	15	1.0	2248	1	V14000	Human p21-activate	678	15	1.0	3440	1	O13967	AGE-modified DNA I
606	15	1.0	2293	1	V59685	Human secreted pro	679	15	1.0	3440	1	V15268	Bce-4 gene with re
607	15	1.0	2309	1	O73786	Human p21-activate	680	15	1.0	3449	1	X52220	Brassica campestris
608	15	1.0	2332	1	V74788	Staphylococcus aur	681	15	1.0	3452	1	V20870	Human phospholipas
609	15	1.0	2334	1	X19798	Human growth reguli	682	15	1.0	3452	1	X13437	Human phospholipas
610	15	1.0	2345	1	V71209	DNA encoding easter	683	15	1.0	3528	1	X13126	Enterococcus faeca
611	15	1.0	2369	1	O77853	Mycoplasma gallise	684	15	1.0	3565	1	O84051	Enterococcus faeca
612	15	1.0	2369	1	V59802	Human secreted pro	685	15	1.0	3589	1	V74576	Sequence encoding
613	15	1.0	2372	1	O49891	Human MDM2 gene. D	686	15	1.0	3589	1	T98531	Staphylococcus aur
614	15	1.0	2372	1	O94589	Human MDM2 gene. N	687	15	1.0	3609	1	V20868	Coding sequence fo
615	15	1.0	2372	1	T45151	Human MDM2-2 gene. C	688	15	1.0	3629	1	T88825	Human Ob-receptor
616	15	1.0	2372	1	T62065	Human MDM2 CDNA. D	689	15	1.0	3689	1	X05815	Mouse pheromone re
617	15	1.0	2372	1	T66410	Human MDM2 coding	690	15	1.0	3721	1	X13348	Enterococcus faeca

C 691	15	1.0	3740	1	Q54682	Potato sucrose pno
C 692	15	1.0	3870	1	T12292	Phospholipase C-ga
C 693	15	1.0	3871	1	T69592	Human Ob receptor
C 694	15	1.0	3893	1	T12293	Phospholipase C-ga
C 695	15	1.0	3896	1	T05816	Mouse pheromone re
C 696	15	1.0	3909	1	T64442	Human haemopoietin
C 697	15	1.0	3921	1	T31300	Rice Xa21 disease
C 698	15	1.0	3972	1	X03040	Human IL-1ra BAC c
C 699	15	1.0	3980	1	V03014	Aspergillus oryzae
C 700	15	1.0	4002	1	Q74463	Endogenous plasmid
C 701	15	1.0	4097	1	V69302	C. elegans OGT DNA
C 702	15	1.0	4102	1	T85575	Human MSX receptor
C 703	15	1.0	4116	1	X02983	Human IL-1ra BAC c
C 704	15	1.0	4118	1	Q29472	Muramidase release
C 705	15	1.0	4403	1	Q73061	SetfABCD gene clust
C 706	15	1.0	4406	1	T74143	Salmonella enterit
C 707	15	1.0	4410	1	Q93913	Yeast MSH1 gene, D
C 708	15	1.0	4434	1	T93389	Bloom's syndrome B
C 709	15	1.0	4437	1	T67013	Bloom's syndrome act
C 710	15	1.0	4437	1	T93392	Bloom's syndrome B
C 711	15	1.0	4437	1	T93390	Bloom's syndrome B
C 712	15	1.0	4437	1	T93394	Bloom's syndrome B
C 713	15	1.0	4437	1	T93395	Bloom's syndrome B
C 714	15	1.0	4438	1	T93391	Bloom's syndrome B
C 715	15	1.0	4438	1	T93393	Bloom's syndrome B
C 716	15	1.0	4454	1	T66462	Plasmid pCRM1.8 co
C 717	15	1.0	4471	1	X00915	Protease activated
C 718	15	1.0	4590	1	N50472	Sequence encoding
C 719	15	1.0	4590	1	T94612	Rat penile neuron
C 720	15	1.0	4779	1	T58751	Nucleolar/endosoma
C 721	15	1.0	4800	1	X03350	Phenol hydroxylase
C 722	15	1.0	4800	1	X01590	Pseudomonas putida
C 723	15	1.0	4818	1	T79909	Actin gene. Promot
C 724	15	1.0	4818	1	T62177	Candida boidinii A
C 725	15	1.0	4879	1	V74400	Staphylococcus aur
C 726	15	1.0	4940	1	Q91286	Orcl gene encoding
C 727	15	1.0	4940	1	T73279	S. cerevisiae orig
C 728	15	1.0	4951	1	X13198	Enterococcus faeca
C 729	15	1.0	5040	1	X20272	Borrelia burgdorfe
C 730	15	1.0	5061	1	V42920	CDNA encoding a hu
C 731	15	1.0	5091	1	Q06164	The lgs gene, enco
C 732	15	1.0	5108	1	Q53403	Sequence encoding
C 733	15	1.0	5232	1	V55038	Human XIAP coding
C 734	15	1.0	5285	1	T80199	Phosphatidyl inosol
C 735	15	1.0	5893	1	V59979	Nucleic acid encod
C 736	15	1.0	5893	1	T75237	Nucleotide sequenc
C 737	15	1.0	6063	1	Q37205	Delta-amino leuall
C 738	15	1.0	6605	1	V23545	Human gravin poly
C 739	15	1.0	6914	1	Q55757	Enterobacter cloac
C 740	15	1.0	7158	1	T27394	Hepatitis E virus
C 741	15	1.0	7168	1	Q45197	HEV strain SAR-55
C 742	15	1.0	7168	1	V71604	Hepatitis E virus
C 743	15	1.0	7194	1	V61690	Hepatitis E virus
C 744	15	1.0	7451	1	V05856	Complete sequence
C 745	15	1.0	7585	1	V68403	Human BAZ2-beta CD
C 746	15	1.0	7997	1	Q89553	Rat cholesterol 7-
C 747	15	1.0	7997	1	T79737	Rat cholesterol 7-
C 748	15	1.0	8982	1	V32452	Human receptor tyr
C 749	15	1.0	9212	1	X13271	Enterococcus faeca
C 750	15	1.0	9444	1	T13279	CDNA to genomic he
C 751	15	1.0	9956	1	V74663	Staphylococcus aur
C 752	15	1.0	10216	1	T39279	Transposon deliver
C 753	15	1.0	10223	1	V52206	Streptococcus pneu
C 754	15	1.0	10726	1	V52199	Streptococcus pneu
C 755	15	1.0	11464	1	V48228	Interleukin 18 con
C 756	15	1.0	11517	1	Q25021	PSF6-SFV4 RNA tran
C 757	15	1.0	11601	1	Q13608	ACV synthetase gen
C 758	15	1.0	13011	1	T96631	CDNA encoding rat
C 759	15	1.0	13058	1	Q48231	Vector containing
C 760	15	1.0	13340	1	X23522	O. longistamnat
C 761	15	1.0	13574	1	X13051	Enterococcus faeca
C 762	15	1.0	13815	1	V18885	Mus musculus dystr
C 763	15	1.0	13999	1	Q43845	Plasmid pAH4611. A
C 764	15	1.0	14231	1	V52214	Streptococcus pneu
C 765	15	1.0	14286	1	X13099	Enterococcus faeca
C 766	15	1.0	15598	1	V74393	Staphylococcus aur
C 767	15	1.0	19307	1	T27558	Shuttle vector pad
C 768	15	1.0	19639	1	X23524	O. longistamnat
C 769	15	1.0	20387	1	X23524	O. longistamnat
C 770	15	1.0	20387	1	V62159	HSV-2 strain SB5 C
C 771	15	1.0	20757	1	X20599	Poly nucleotide seq
C 772	15	1.0	26338	1	V62134	HSV-2 strain SB5 C
C 773	15	1.0	28994	1	V15826	Genomic DNA for in
C 774	15	1.0	29729	1	X13175	Enterococcus faeca
C 775	15	1.0	32768	1	X13037	Enterococcus faeca
C 776	15	1.0	34446	1	X24332	Bovine adenovirus
C 777	15	1.0	80595	1	V83939	HC-contig derived
C 778	15	1.0	110000	1	T58840_2	Mycoplasma genital
C 779	15	1.0	110000	1	T58840_3	Continuation (3 of
C 780	15	1.0	110000	1	T58840_4	Continuation (4 of
C 781	15	1.0	110000	1	T58840_5	Continuation (5 of
C 782	15	1.0	110000	1	V30458_1	Continuation (2 of
C 783	15	1.0	110000	1	V30458_3	Continuation (2 of
C 784	15	1.0	110000	1	V21209_01	Continuation (2 of
C 785	15	1.0	110000	1	V21209_02	Continuation (3 of
C 786	15	1.0	110000	1	V21209_03	Continuation (3 of
C 787	15	1.0	110000	1	V21209_04	Continuation (3 of
C 788	15	1.0	110000	1	V21209_05	Continuation (5 of
C 789	15	1.0	110000	1	V21209_06	Continuation (6 of
C 790	15	1.0	110000	1	V21209_08	Continuation (9 of
C 791	15	1.0	110000	1	V21209_13	Continuation (14 of
C 792	15	1.0	110000	1	V21209_14	Continuation (15 of
C 793	15	1.0	110000	1	V21209_15	Continuation (15 of
C 794	15	1.0	110000	1	X20248_07	Continuation (8 of
C 795	15	1.0	110000	1	X20248_08	Continuation (9 of
C 796	15	1.0	110000	1	V30459_1	Continuation (2 of
C 797	15	1.0	117213	1	V30459_3	Continuation (4 of
C 798	15	1.0	117213	1	V62176	Continuation (4 of
C 799	15	1.0	237326	1	V57903	HSV-2 strain SB5 C
C 800	15	0.9	237326	1	T57389	Hereditary haemoch
C 801	15	0.9	237326	1	T57389	RSV N haemoch
C 802	15	0.9	237326	1	T57389	CD4 gene exon 9a
C 803	15	0.9	237326	1	T38259	Y. pestis V antigen
C 804	15	0.9	237326	1	V03040	Sheep beta globin
C 805	15	0.9	237326	1	V82050	S. pneumoniae hist
C 806	15	0.9	237326	1	Q72562	Regulatory region
C 807	15	0.9	237326	1	T91947	Methanol regulated
C 808	15	0.9	237326	1	T91947	ORF3 derived cance
C 809	15	0.9	237326	1	T91947	PCR primer 2 used
C 810	15	0.9	237326	1	V52820	Antiense PCR prim
C 811	15	0.9	237326	1	T64665	G-CSF receptor ago
C 812	15	0.9	237326	1	V63083	Human ADP-ribosylt
C 813	15	0.9	237326	1	V71173	Primer for pUC17I
C 814	15	0.9	237326	1	T58233	Bacillus thuringie
C 815	15	0.9	237326	1	T66493	Primer for pUC17I
C 816	15	0.9	237326	1	V41603	Fucosyl isomerase C
C 817	15	0.9	237326	1	V41603	Nucleotide sequenc
C 818	15	0.9	237326	1	V41603	Nucleotide sequenc
C 819	15	0.9	237326	1	Q58217	Synthetic FLAG seq
C 820	15	0.9	237326	1	T64670	G-CSF receptor ago
C 821	15	0.9	237326	1	V41606	Nucleotide sequenc
C 822	15	0.9	237326	1	T29648	Primer for amplify
C 823	15	0.9	237326	1	T51219	Primer #1 for huma
C 824	15	0.9	237326	1	T68712	PCR primer for huma
C 825	15	0.9	237326	1	T54076	Human hcp SH2 doma
C 826	15	0.9	237326	1	V58134	Human hcp SH2 doma
C 827	15	0.9	237326	1	T37295	Human Src primer #
C 828	15	0.9	237326	1	T51213	Primer #1 for huma
C 829	15	0.9	237326	1	T51217	Primer #1 for huma
C 830	15	0.9	237326	1	T51217	Primer #1 for huma
C 831	15	0.9	237326	1	T51217	Primer #1 for huma
C 832	15	0.9	237326	1	T51217	Primer #1 for huma
C 833	15	0.9	237326	1	T51217	Primer #1 for huma
C 834	15	0.9	237326	1	T51217	Primer #1 for huma
C 835	15	0.9	237326	1	T51217	Primer #1 for huma
C 836	15	0.9	237326	1	V49400	Primer AB177 for E

C 837	14	0.9	44	1	V49404	Primer AB181 for W	C 910	14	0.9	315	1	T26332	Human gene signatu
C 838	14	0.9	44	1	V49406	Primer AB179 for W	C 911	14	0.9	318	1	O99373	Rat allograft infl
C 839	14	0.9	44	1	V49408	Primer AB185 for V	C 912	14	0.9	319	1	V89670	EST clone DA244. N
C 840	14	0.9	45	1	T36190	5' PCR primer for	C 913	14	0.9	325	1	V78583	Staphylococcus aur
C 841	14	0.9	45	1	T37297	Human lck SH2 doma	C 914	14	0.9	327	1	V77350	Staphylococcus aur
C 842	14	0.9	45	1	T51215	Primer #1 for huma	C 915	14	0.9	331	1	O60313	Human brain expres
C 843	14	0.9	45	1	T68708	PCR primer for hum	C 916	14	0.9	336	1	V80245	Modified Brazil nu
C 844	14	0.9	45	1	V01771	C. trachomatis cry	C 917	14	0.9	337	1	V77399	Staphylococcus aur
C 845	14	0.9	45	1	V49410	Primer AB183 for V	C 918	14	0.9	342	1	O72503	Dermatophagoides f
C 846	14	0.9	45	1	V68583	Nucleotide sequenc	C 919	14	0.9	351	1	O59508	Human brain expres
C 847	14	0.9	45	1	V82661	Amplification prim	C 920	14	0.9	359	1	T26802	Human gene signatu
C 848	14	0.9	45	1	X30265	Chlamydia cryptic	C 921	14	0.9	362	1	V75990	Staphylococcus aur
C 849	14	0.9	47	1	X30006	Amplification prim	C 922	14	0.9	367	1	T22133	Human gene signatu
C 850	14	0.9	47	1	V49441	Primer AB111 for B	C 923	14	0.9	367	1	V80248	Human gene signatu
C 851	14	0.9	49	1	T04740	PCR primer, NCCPCR	C 924	14	0.9	367	1	V80249	Modified Brazil nu
C 852	14	0.9	56	1	O69027	Kanamycin resistan	C 925	14	0.9	367	1	V80250	Modified Brazil nu
C 853	14	0.9	56	1	O78549	Zipper adapter RAQ	C 926	14	0.9	367	1	V80251	Modified Brazil nu
C 854	14	0.9	56	1	T80838	Gene specific zip	C 927	14	0.9	367	1	V80247	Modified Brazil nu
C 855	14	0.9	63	1	O56901	Encodes enterokina	C 928	14	0.9	368	1	O58788	Human brain expres
C 856	14	0.9	99	1	M61372	Synthetic regulati	C 929	14	0.9	368	1	V78587	Staphylococcus aur
C 857	14	0.9	100	1	M61373	Synthetic regulati	C 930	14	0.9	370	1	O59401	Human brain expres
C 858	14	0.9	101	1	M61370	Synthetic regulati	C 931	14	0.9	371	1	V78508	Staphylococcus aur
C 859	14	0.9	101	1	M61371	Synthetic regulati	C 932	14	0.9	374	1	O39922	Expressed Sequence
C 860	14	0.9	102	1	O76487	Human genome fragm	C 933	14	0.9	374	1	O59334	Human brain expres
C 861	14	0.9	105	1	V69317	Mycobacterium sp.	C 934	14	0.9	374	1	X20698	Poly nucleotide seq
C 862	14	0.9	117	1	N71019	Sequence of fusion	C 935	14	0.9	376	1	T19954	Human gene signatu
C 863	14	0.9	149	1	V77203	Staphylococcus aur	C 936	14	0.9	382	1	T19522	Human gene signatu
C 864	14	0.9	150	1	X11800	Human biallelic po	C 937	14	0.9	390	1	T09150	S. haematobium trv
C 865	14	0.9	150	1	X11858	Human biallelic po	C 938	14	0.9	399	1	V89560	EST clone CR422. N
C 866	14	0.9	150	1	X10406	Human biallelic po	C 939	14	0.9	400	1	O85036	DNA-ase-B2 gene fr
C 867	14	0.9	154	1	X10388	Human biallelic po	C 940	14	0.9	400	1	T12773	S. pyogenes DNasee
C 868	14	0.9	156	1	X30593	H. pylori inner me	C 941	14	0.9	400	1	V78199	Staphylococcus aur
C 869	14	0.9	159	1	T18603	Human chondrocyte	C 942	14	0.9	400	1	V78165	Staphylococcus aur
C 870	14	0.9	175	1	X11708	Human biallelic po	C 943	14	0.9	400	1	V78031	Staphy

Query Match	Best Local Similarity	Score 927	DB 1	Length 1014
Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				

ALIGNMENTS

RESULT 1

ID T04222 standard; DNA; 1014 BP.

AC T04222;

DT 18-APR-1996 (first entry)

DE Partial lcrV (V antigen) gene of *Y. pestis*.

KW lcrV; V antigen; virulence; plague; vaccine; epitope; ss.

OS *Yersinia pestis*.

FT Key Location/Qualifiers

FT cds 1..990

FT /note- "V antigen"

FT /tag- a

FT W09524475-A1.

PD 14-SEP-1995.

PF 06-MAR-1995; G00481.

PR 08-MAR-1994; GB-004577.

PA (MINA) UK SEC FOR DEFENCE.

PI Leary SEC, Tibball RW, Williamson ED, Leary SE.

DR W01, 95-328268/42.

DR P-PSDB; R79961.

PT Recombinant DNA expressing *Yersinia pestis* V antigen - useful in oral or parenteral vaccines for protection against plague

PS Claim 6; Page 11-13; 25pp; English.

CC T04222-23 are DNA sequences (lcrV) encoding all or a protective epitope part of the mature V protein of *Yersinia pestis*. The protein was expressed as a fusion protein with maltose binding protein or glutathione-S-transferase in 3 different plasmid vectors. *Y. pestis* is the highly virulent causative organism of plague in a wide range of animals, including man. The V antigen (lcrV) is an unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen is postulated to act as a virulence antigen, and transformed microorganisms cony. recombinant DNA encoding a V antigen protein/peptide are useful in vaccines to protect against plague.

CC Sequence 1014 BP; 346 A; 181 C; 201 G; 286 T;

Query Match 59.28; Score 927; DB 1; Length 1014;

Best Local Similarity 99.98; Pred. No. 0;

Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 589 ATTGAGGCTCGAACAACCAACCAACATTTATTTGAGTCTAGAAAAGTAGGCTG 648

DB 13 ATTGAGGCTCGAACAACCAACCAACATTTATTTGAGTCTAGAAAAGTAGGCTG 72

QY 649 GACAACTTACTGTCATGCTTCTTCATTTAGAGATTTGCTCAAGTCAAGAT 708

DB 73 GACAACTTACTGTCATGCTTCTTCATTTAGAGATTTGCTCAAGTCAAGAT 132

QY 709 AAAATATAGATATTTTCATTAATATGATCCGAAAAGTTGGAGGTTTGGCCAT 768

Query Match	Best Local Similarity	Score 927	DB 1	Length 1014
Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				

RESULT 2

ID T38243 standard; DNA; 1014 BP.

AC T38243;

DT 28-DEC-1996 (first entry)

DE *Y. pestis* lcrV (V antigen) gene.

KW Plague; vaccine; genetic immunisation; V antigen; lcrV;

OS *Yersinia pestis* strain GB.

FT Key Location/Qualifiers

FT cds 1..990

FT /tag- a

FT /tag- b

FT misc_feature 1..10

CC PMAL-c2 or pGEX-5x-2 (see also T38243) to allow prodn. of
 CC recombinant V antigen for use in vaccines against plague.
 CC Expression in gut-colonizing organisms and attenuated Salmonella
 CC typhl allows prodn. of live vaccines. Fl/V antigen fusions were
 CC also created (see also T38249 and T38256). The gene can itself be
 CC used in genetic vaccines
 SQ Sequence 1014 BP; 347 A; 180 C; 201 G; 286 T;

Query Match 55.9%; Score 876; DB 1; Length 1014;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 976; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 589 ATTAGAGCCTACGACAAACCCACACATTTATTGAGGATCTAGAAAAAGTAGGGTG 648
 DB 13 ATTAGAGCCTACGACAAACCCACACATTTATTGAGGATCTAGAAAAAGTAGGGTG 72
 QY 649 GAACAACCTTACTGTCATGCTTCTTCAGTTTGAAGAATGGTTCACTTACCAAGAT 708
 DB 73 GAACAACCTTACTGTCATGCTTCTTCAGTTTGAAGAATGGTTCACTTACCAAGAT 132
 QY 709 AAAAATATAGATATTTCCATTAAATATATGATCCAGAAAAAGATTGGAGTTTGGCAAT 768
 DB 133 AAAAATATAGATATTTCCATTAAATATATGATCCAGAAAAAGATTGGAGTTTGGCAAT 192
 QY 769 AGAGTAATTACTGATGATATGCAATGCTCAGAGAAATCCCTAGCTTATTTTCCACCGAG 828
 DB 193 AGAGTAATTACTGATGATATGCAATGCTCAGAGAAATCCCTAGCTTATTTTCCACCGAG 252
 QY 829 GATACCATCTTAAAGCGGCTCATTTATGACAAACCACTGCAAAATGGCATCAAGCGAGTA 888
 DB 253 GATGCACTCTTAAAGCGGCTCATTTATGACAAACCACTGCAAAATGGCAACGAGGTA 312
 QY 889 AAAGAGTTCCTTGAATCATGCGCCGATACCAATGGAATGGGAGGCTTCATGGCAGTA 948
 DB 313 AAAGAGTTCCTTGAATCATGCGCCGATACCAATGGAATGGGAGGCTTCATGGCAGTA 372
 QY 949 ATGCATTTCTCTTAAACCGCGATCGTATGATGATATTTTGAAGTATGTTGAT 1008
 DB 373 ATGCATTTCTCTTAAACCGCGATCGTATGATGATATTTTGAAGTATGTTGAT 432
 QY 1009 TCATATATCATATGATGATGCGCGTAGCAATTCGCGTAAATATAGTGTAGCTTACC 1068
 DB 433 TCATATATCATATGATGATGCGCGTAGCAATTCGCGTAAATATAGTGTAGCTTACC 492
 QY 1069 GCCGATTTAAAGATTTATTCAGTTTCAAGCCGAATTAATAGATCGTGTAGTACT 1128
 DB 493 GCCGATTTAAAGATTTATTCAGTTTCAAGCCGAATTAATAGATCGTGTAGTACT 552
 QY 1129 GCCGATTTAAATATTCATGATTAATCCATTATCTCATGATGATAAATTTATATGTTAT 1188
 DB 553 GCCGATTTAAATATTCATGATTAATCCATTATCTCATGATGATAAATTTATATGTTAT 612
 QY 1189 ACAGATGAGAGATTTTAAAGCCAGCGAGATACAAAATTTCTCGAAGAAATGCTCAA 1248
 DB 613 ACAGATGAGAGATTTTAAAGCCAGCGAGATACAAAATTTCTCGAAGAAATGCTCAA 672
 QY 1249 ACCACATTCAGTGTGATGGAGCGAGAAAAAATAGTCTCGATTAAGAGCTTCTTGGG 1308
 DB 673 ACCACATTCAGTGTGATGGAGCGAGAAAAAATAGTCTCGATTAAGAGCTTCTTGGG 732
 QY 1309 AGTGAGATTAAGAAACCGGGGGCTTGGGTAATCTGAAAAAATCAATCTTTATATAA 1368
 DB 733 AGTGAGATTAAGAAACCGGGGGCTTGGGTAATCTGAAAAAATCAATCTTTATATAA 792
 QY 1369 GATTAATATGATATTTCTCTTTCGACACACTGCTCGGATTAAGTCCAGCGCTCAAC 1428
 DB 793 GATTAATATGATATTTCTCTTTCGACACACTGCTCGGATTAAGTCCAGCGCTCAAC 852
 QY 1429 GACTTGTGAGCAAAAAACACTAGCTGTCTGATATTTATCATCAGGTTTATATAGCT 1488
 DB 853 GACTTGTGAGCAAAAAACACTAGCTGTCTGATATTTATCATCAGGTTTATATAGCT 912

QY 1489 ATTGAAGCACTGACACCGTTTCATTCAGAAATATGATTCAGTATGCAACGTCTGTAGAT 1548
 DB 913 ATTGAAGCACTGACACCGTTTCATTCAGAAATATGATTCAGTATGCAACGTCTGTAGAT 972
 QY 1549 GACACGCTCTGCTTAAGA 1566
 DB 973 GACACGCTCTGCTTAAGA 990

RESULT 4

T38256
 ID T38256 standard; DNA; 1462 BP.
 AC T38256;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis Fl/V antigen gene fusion.
 KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
 KW Fl antigen; catf; ds.
 OS Chimeric Yersinia pestis strain GB;
 OS Chimeric synthetic.
 FH Key Location/Qualifiers
 FT cds 8..1450
 FT /tag- a
 FT /product- Fl/V fusion protein
 FT 452..472
 FT /note- b
 FT /note- "bases 452-472 is a sequence derived
 from PCR primers"

W09628551-1.

PD 19-SEP-1996
 PF 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
 DR WPI; 96-433824/43.
 P-PSDB; W01045.
 PT Yersinia pestis V antigen and Fl antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS disclosure; Page 51-55; 98pp; English.
 CC A gene fusion (T38256) comprises coding sequences for the Yersinia
 CC pestis Fl antigen (see also T38244) (without the signal sequence)
 CC and for the Y. pestis V antigen (see also T38243), joined by a
 CC sequence encoding a 6-amino acid peptide linker. It was obtd. by
 CC PCR amplification of the individual genes using primers (see also
 CC T38245) based on the genes and including linker sequences. The
 CC gene fusion (see also T38249) can be used to produce Fl/V fusion
 CC protein (W01045) in transformed cells, esp. gut-colonising
 CC organisms, to induce an immune response against Y. pestis, the
 CC causative organism of plague.
 SQ Sequence 1462 BP; 476 A; 285 C; 301 G; 400 T;

Query Match 55.9%; Score 876; DB 1; Length 1462;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 976; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 589 ATTAGAGCCTACGACAAACCCACACATTTATTGAGGATCTAGAAAAAGTAGGGTG 648
 DB 473 ATTAGAGCCTACGACAAACCCACACATTTATTGAGGATCTAGAAAAAGTAGGGTG 532
 QY 649 GAACAACCTTACTGTCATGCTTCTTCAGTTTGAAGAATGGTTCACTTACCAAGAT 708
 DB 533 GAACAACCTTACTGTCATGCTTCTTCAGTTTGAAGAATGGTTCACTTACCAAGAT 592
 QY 709 AAAAATATAGATATTTCCATTAAATATGATCCAGAAAAAGATTGGAGTTTGGCAAT 768
 DB 593 AAAAATATAGATATTTCCATTAAATATGATCCAGAAAAAGATTGGAGTTTGGCAAT 652
 QY 769 AGAGTAATTACTGATGATATGCAATTTGCTCAGAAAAATCCTGTTATTTTACCCGAG 828
 DB 653 AGAGTAATTACTGATGATATGCAATTTGCTCAGAAAAATCCTGTTATTTTACCCGAG 712

Accession	Source	Length	Strain	Notes
QY 1189	ACAGATGAAAGATTTTAAAGCCAGCGAGAGTACAAAATTCGAGAAAATGCTTCA	1248		
Db 1141	ACAGATGAAAGATTTTAAAGCCAGCGAGAGTACAAAATTCGAGAAAATGCTTCA	1200		
QY 1249	ACCACATTACAGTGGATGGAGCGAGAAAAAATAGTCTCGATAAAGACTTCTTGA	1308		
Db 1201	ACCACATTACAGTGGATGGAGCGAGAAAAAATAGTCTCGATAAAGACTTCTTGA	1260		
QY 1309	AGTGAATATAAAGAACCGGGCGTTGGGTAAATCTGAAAAACCTACATCTATATAAA	1368		
Db 1261	AGTGAATATAAAGAACCGGGCGTTGGGTAAATCTGAAAAACCTACATCTATATAAA	1320		
QY 1369	GATTAATTAATGAATTTTCTCTACTCTTTGGCCACACCTGCTCGATAAGTCAGCGCTCAAC	1428		
Db 1321	GATTAATTAATGAATTTTCTCTACTCTTTGGCCACACCTGCTCGATAAGTCAGCGCTCAAC	1380		
QY 1429	GACTTGTTAGCCAAAAACAACCTACAGCTGTCTGATATTACATCAGCTTTTAATTCAGCT	1488		
Db 1381	GACTTGTTAGCCAAAAACAACCTACAGCTGTCTGATATTACATCAGCTTTTAATTCAGCT	1440		
QY 1489	ATTGAAGCCCTACACCGTTTCATTTCGAAATTAATGATTCAGTATGCAAGCTCGTAGAT	1548		
Db 1441	ATTGAAGCCCTACACCGTTTCATTTCGAAATTAATGATTCAGTATGCAAGCTCGTAGAT	1500		
QY 1549	GACACGCTCTGTAAATGA 1566			
Db 1501	GACACGCTCTGTAAATGA 1518			

RESULT 6

Accession	Source	Length	Strain	Notes
T04223	T04223 standard; DNA; 1014 BP.			
ID 104223				
AC T04223				
DE 18-APR-1996	(first entry)			
DT Partial LcrV (V antigen) gene of Y. pestis.				
KM LcrV; V antigen; virulence; plague; vaccine; epitope; ss.				
OS Yersinia pestis.				
OS Key	Location/Qualifiers			
FT cds	1..990			
FT	/*tag= a			
FN MO9524475-A1.				
PD 14-SEP-1995.				
PF 06-MAR-1995; G00481.				
PR 08-MAR-1994; GB-004577.				
PA (MINA) UK SEC FOR DEFENCE.				
PI Leary SEC, Titball RW, Williamson ED, Leary SE;				
DR WPI; 95-328268/42.				
DR P-SDS; R79962.				
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in				
PT oral or parenteral vaccines for protection against plague				
PS Claim 6; Page 15-16; 25pp; English.				
CC T04222-23 are DNA sequences (LcrV) encoding all or a protective epitopic				
CC part of the mature V protein of Yersinia pestis. The protein was				
CC expressed as a fusion protein with maltose binding protein or				
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis				
CC is the highly virulent causative organism of plague in a wide range of				
CC animals, including man. The V antigen (LcrV) is an unstable 37.3 kDa				
CC monomeric peptide encoded on the ca. 70 kb Lcr plasmid. The V antigen				
CC is postulated to act as a virulence antigen, and transformed				
CC microorganisms contg. recombinant DNA encoding a V antigen protein/				
CC peptide are useful in vaccines to protect against plague.				
CC Sequence 1014 BP; 343 A; 185 C; 204 G; 282 T.				

Query Match	55.78;	Score 872;	DB 1;	Length 1014;
Best Local Similarity	99.88;	Pred. No. 0;		
Matches 972;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

[illegible]

Db	77	AACCTACGGC	CANGGTTCTTCAGTTTAGAAGAA	TGGTTACGTAGCAAGATATAA	136
Oy	713	ATATAGATATTT	CCATTAAATATGATCCAGAAA	GATTGAGGTTTTGGCAATAGAG	772
Db	137	ATATAGATATTT	CCATTAAATATGATCCAGAAA	GATTGAGGTTTTGGCAATAGAG	196
Oy	773	TAATTACGATG	ATATGATTCGTAAGAAA	TCTACTTAATTTTCACCCGSGATA	832
Db	197	TAATTACGATG	ATATGATTCGTAAGAAA	TCTACTTAATTTTCACCCGSGATG	256
Oy	833	CCATTCTTAAAG	GCGGTCATTATGACAA	CCAACTGCAAAATGGCATCAACGAGTAAAG	892
Db	257	CCATTCTTAAAG	GCGGTCATTATGACAA	CCAACTGCAAAATGGCATCAACGAGTAAAG	316
Oy	893	AGTTCTTTGAAT	CATCGCCGAATACACAA	TGGGAATGGCGGCGTTTCATGGCAGTAATGC	952
Db	317	AGTTCTTTGAAT	CATCGCCGAATACACAA	TGGGAATGGCGGCGTTTCATGGCAGTAATGC	376
Oy	953	ATTCTCTTTAAC	CGCGGATCGATGATGATGATATTTT	TGAAAAGTATGTTGATTCAA	1012
Db	377	ATTCTCTTTAAC	CGCGGATCGATGATGATGATATTTT	TGAAAAGTATGTTGATTCAA	436
Oy	1013	TGAATCATCAT	GATGTCATCCCGTAGCAAGTTGGCGTA	GAAATTAGCTGAGCTTACCGCGG	1072
Db	437	TGAATCATCAT	GATGTCATCCCGTAGCAAGTTGGCGTA	GAAATTAGCTGAGCTTACCGCGG	496
Oy	1073	AATTAAGATTTT	ATTCAGTTATTCACAGCCGAATTA	ATAGATCTGCTCTAGTACTGCA	1132
Db	497	AATTAAGATTTT	ATTCAGTTATTCACAGCCGAATTA	ATAGATCTGCTCTAGTACTGCA	556
Oy	1133	CCATTAATATTC	ACAGATAAATCCATTATCTCATGAT	AAAAATTTATATGTTATACAG	1192
Db	557	CCATTAATATTC	ACAGATAAATCCATTATCTCATGAT	AAAAATTTATATGTTATACAG	616
Oy	1193	ATGAAGAGATTTT	AAAAAGCCAGCGCAGAGTACAAA	ATTCGAGAAAATGCTCTAAACCA	1252
Db	617	ATGAAGAGATTTT	AAAAAGCCAGCGCAGAGTACAAA	ATTCGAGAAAATGCTCTAAACCA	676
Oy	1253	CCATTACAGG	TGGGATGGGAGGAGAAAAA	ATGTCGATAAGGACTTCTTGGAAGTG	1312
Db	677	CCATTACAGG	TGGGATGGGAGGAGAAAAA	ATGTCGATAAGGACTTCTTGGAAGTG	736
Oy	1313	AGAAATAAAGAA	CCGGGCGTTGGGTATCTGAAAAA	ACTCATCTTATATAATAAGATA	1372
Db	737	AGAAATAAAGAA	CCGGGCGTTGGGTATCTGAAAAA	ACTCATCTTATATAATAAGATA	796
Oy	1373	ATAATGATTTT	CTCACTTTGGCCACCA	CCCTGCTCGGATATGTCACAGCGCTCAACGACT	1432
Db	797	ATAATGATTTT	CTCACTTTGGCCACCA	CCCTGCTCGGATATGTCACAGCGCTCAACGACT	856
Oy	1433	TGGTTAGCAAAA	AAACAACTACAGCTGTCGATATTA	TCATCAGAGTTTATATAGCTATG	1492
Db	857	TGGTTAGCAAAA	AAACAACTACAGCTGTCGATATTA	TCATCAGAGTTTATATAGCTATG	916
Oy	1493	AAGCACTGAAC	CGTTTCATTACAGAAATATGATTC	AGTATGCAACGTCGCTAGATGACA	1552
Db	917	AAGCACTGAAC	CGTTTCATTACAGAAATATGATTC	AGTATGCAACGTCGCTAGATGACA	976
Oy	1553	CGTCTGGTAA	TATGA	1566	
Db	977	CGTCTGGTAA	TATGA	990	

RESULT	7
T38248	
ID	T38248 standard; DNA; 547 bp.
AC	T38248;
DT	28-Dec-1996 (first entry)
DE	Y. pestis F1 antigen catl gene (including signal sequence).
KW	Plague; vaccine; genetic immunisation; F1 antigen; catl;
KW	Y antigen; ds.
OS	Yersinia pestis strain GB.

Key Location/Qualifiers
 signal_peptide 29..91
 FT /tag= a
 FT mat_peptide 92..538
 FT /*tag= b
 W09628551-A1.
 PD 19-SEP-1996.
 PF 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (M1NA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tlball RM, Williamson ED;
 DR WPI: 96-433824/43.
 DR P-PSDB: W01043.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Clam 41; Page 61-62; 96pp; English.
 CC A DNA sequence (T38248) comprises the cafI gene, including the
 CC signal sequence, coding for the F1 antigen (W01043) of Yersinia
 CC pestis. It was obt'd. by PCR amplification (see also T38257-58)
 CC of Y. pestis DNA. The PCR product was cloned into plasmid pKCMV
 CC and the resulting plasmid (pF1AB) was used to transform E. coli Nova
 CC Blue. Purified plasmid, when administered by i.m. injection,
 CC induced an immunoglobulin response to F1 in BALB/c mice. Live
 CC vaccines comprising gut colonising organisms transformed with the
 CC cafI gene (see also T38244) can be used to protect a host animal
 CC against plague.
 SQ Sequence 547 BP; 165 A; 120 C; 115 G; 147 T;

Query Match 32.8%; Score 513; DB 1; Length 547;
 Best Local Similarity 100.0%; Pred. No. 1.4e-249;
 Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 68 ATATGAAAAAATCAGTCCGTTATCGCATTCATTTGGAACATTCGACTA 127
 DB 27 ATATGAAAAAATCAGTCCGTTATCGCATTCATTTGGAACATTCGACTA 86
 0Y 128 ATGGCGGAGATTACTGCAACGACACGACGCACTCTGTGTAACGAGCCGCA 187
 DB 87 ATGGCGGAGATTACTGCAACGACACGACGCACTCTGTGTAACGAGCCGCA 146
 0Y 188 TCACCTTACATATAAGAAAGCGGCTCAATATGCAATTTGCAATGAAACATGATA 247
 DB 147 TCACCTTACATATAAGAAAGCGGCTCAATATGCAATTTGCAATGAAACATGATA 206
 0Y 248 CAGAAATTACTTGTGGTACGCTTACTCTTGGGGCTATTAACAGAACACATGACAT 307
 DB 207 CAGAAATTACTTGTGGTACGCTTACTCTTGGGGCTATTAACAGAACACATGACAT 266
 0Y 308 CTGTAACTTACATAGATGCGGCGGCTGATCCATGATTAACATTTACTCTCAGATG 367
 DB 267 CTGTAACTTACATAGATGCGGCGGCTGATCCATGATTAACATTTACTCTCAGATG 326
 0Y 368 GAAATACCAACCAATTCATCAAAAAGTATGGCAAGATTCAGATTTTGATATCT 427
 DB 327 GAAATACCAACCAATTCATCAAAAAGTATGGCAAGATTCAGATTTTGATATCT 386
 0Y 428 CTCCTAAGTAAAGCGTGAAGACCTTGTGGGGATGACGCTGCTTGGCTAGGCGGAC 487
 DB 387 CTCCTAAGTAAAGCGTGAAGACCTTGTGGGGATGACGCTGCTTGGCTAGGCGGAC 446
 0Y 488 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGCTAACTGACAGAGTAATATA 547
 DB 447 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGCTAACTGACAGAGTAATATA 506
 0Y 548 CTGATGCTTAACCGTAAACGATATCAACCAAG 580
 DB 507 CTGATGCTTAACCGTAAACGATATCAACCAAG 539

RESULT 8

092819
 ID 092819 standard; DNA: 542 BP.
 AC 092819;
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis cafI (F1) antigen in plasmid pFORF1B.
 KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 KM bubonic plague; pneumonic plague; ds.
 OS Yersinia pestis.
 FH Key
 FT Location/Qualifiers
 FT cds 2..7
 FT /*tag= a
 FT /note= "first protein encoded by pFORF1B"
 FT misc_feature 1..6
 FT /*tag= b
 FT misc_feature 536..541
 FT /*tag= c
 FT /note= "cafI open reading frame downstream seq."
 FT /*tag= d
 FT /note= "cafI fusion protein"
 PN W09518231-A1.
 PD 06-JUL-1995.
 PF 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (M1NA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tlball RM, Williamson ED;
 DR WPI: 95-246396/32.
 DR P-PSDB: R76528.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure: Page 19-20; 27pp; English.
 CC The sequence represents the plasmid pFORF1B including the entire
 CC Y. pestis cafI (F1) antigen gene having a 5' tail including a SacI
 CC restriction site, and up to TATAG downstream of the cafI ORF.
 CC The DNA construct can be used to transform human or animal gut
 CC colonizing microorganisms, specifically attenuated Salmonella
 CC typhimurium or Salmonella typhi. The transformed microorganisms
 CC can be used as live/attenuated vaccines which induce immune
 CC responses at mucosal surfaces. The vaccines provide protection
 CC against infection with Y. pestis, and are parenterally and orally
 CC active vaccines offering protection against bubonic and pneumonic
 CC plague.
 SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

Query Match 32.7%; Score 512; DB 1; Length 542;
 Best Local Similarity 100.0%; Pred. No. 4.4e-249;
 Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 68 ATATGAAAAAATCAGTCCGTTATCGCATTCATTTGGAACATTCGACTGCTA 127
 DB 19 ATATGAAAAAATCAGTCCGTTATCGCATTCATTTGGAACATTCGACTGCTA 78
 0Y 128 ATGGCGGAGATTACTGCAACGACACGACGCAACTCTTGTAAACAGGCCCA 187
 DB 79 ATGGCGGAGATTACTGCAACGACACGACGCAACTCTTGTAAACAGGCCCA 138
 0Y 188 TCACCTTACATATAAGAAAGCGGCTCAATATGCAATTTGGAACATTCGACTGATA 247
 DB 139 TCACCTTACATATAAGAAAGCGGCTCAATATGCAATTTGGAACATTCGACTGATA 198
 0Y 248 CAGAAATTACTTGTGGTACGCTTACTCTTGGCGGCTATTAACAGAAACCACTAGACAT 307
 DB 199 CAGAAATTACTTGTGGTACGCTTACTCTTGGCGGCTATTAACAGAAACCACTAGACAT 258
 0Y 308 CTGTAACTTACATAGATGCGGCGGCTGATCCATGATTAACATTTACTCTCAGATG 367
 DB 259 CTGTAACTTACATAGATGCGGCGGCTGATCCATGATTAACATTTACTCTCAGATG 318
 0Y 368 GAAATACCAACCAATTCATCAAAAAGTATGGCAAGATTCAGATTTTGATATCT 427
 DB 319 GAAATACCAACCAATTCATCAAAAAGTATGGCAAGATTCAGATTTTGATATCT 378

QY 428 CTCCTAAGTAAACGGTGAAGACCTTGTGGGGATGACGTCTTGTGGCTACGGGACGCC 487
DB 379 CTCCTAAGTAAACGGTGAAGACCTTGTGGGGATGACGTCTTGTGGCTACGGGACGCC 438
QY 488 AGAATTTCTTGTTCCTCAATTTGGTTCCAAAGCGCGTAACTTGCAGCAGTAATAACA 547
DB 439 AGAATTTCTTGTTCCTCAATTTGGTTCCAAAGCGCGTAACTTGCAGCAGTAATAACA 498
QY 548 CTGATGCTGAACCGTAAACCGTATCTAACCA 579
DB 499 CTGATGCTGAACCGTAAACCGTATCTAACCA 530

RESULT 9

V41594
ID V41594 standard: DNA; 544 Bp.
AC V41594:
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nyp1(a)sec544.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 17..529
FT /tag= a
FT /product= "F1 antigen"
PN MO9824912-A2.
PD 11-JUN-1998.
PE 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR MPI; 98-33331/29.
DR P-PSDB; W59782.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 51; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 544 Bp; 164 A; 120 C; 114 G; 146 T;

Query Match 32.7%; Score 512; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 4.4e-249;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCACTCCGTTATCGCATTCATTTGGAACTATTGCAACTGCTA 127
DB 15 ATATGAAAAAATCACTCCGTTATCGCATTCATTTGGAACTATTGCAACTGCTA 74
QY 128 ATGGGGAGATTAACTGCAAGCAGCACTGCAACGGCACTCTGTGAACGCCGCCA 187
DB 75 ATGGGGAGATTAACTGCAAGCAGCACTGCAACGGCACTCTGTGAACGCCGCCA 134
QY 188 TCACCTTACATATAAGAGAGCGCTCCATTAATTAATGAGCAATGAGAACTGATA 247
DB 135 TCACCTTACATATAAGAGAGCGCTCCATTAATTAATGAGCAATGAGAACTGATA 194
QY 248 CAGAAATTAATGTTGGTGAAGCTTACTTGTGGCGCTATAAAGCAGAACACTAGCAT 307
DB 195 CAGAAATTAATGTTGGTGAAGCTTACTTGTGGCGCTATAAAGCAGAACACTAGCAT 254
QY 308 CTGTTAATTAAGATGCGCGGGGTATGCCATGTAATTAATTAATTAATTAATTA 367
DB 255 CTGTTAATTAAGATGCGCGGGGTATGCCATGTAATTAATTAATTAATTAATTA 314
QY 368 GAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 427
DB 315 GAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 374

QY 428 CTCCTAAGTAAACGGTGAAGACCTTGTGGGGATGACGTCTTGTGGCTACGGGACGCC 487
DB 375 CTCCTAAGTAAACGGTGAAGACCTTGTGGGGATGACGTCTTGTGGCTACGGGACGCC 434
QY 488 AGAATTTCTTGTTCCTCAATTTGGTTCCAAAGCGCGTAACTTGCAGCAGTAATAACA 547
DB 435 AGAATTTCTTGTTCCTCAATTTGGTTCCAAAGCGCGTAACTTGCAGCAGTAATAACA 494
QY 548 CTGATGCTGAACCGTAAACCGTATCTAACCA 579
DB 495 CTGATGCTGAACCGTAAACCGTATCTAACCA 526

RESULT 10

V41596
ID V41596 standard: DNA; 544 Bp.
AC V41596:
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nyp1sec510.
KW F1 antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 17..532
FT /tag= a
FT /product= "F1 antigen"
PN MO9824912-A2.
PD 11-JUN-1998.
PE 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR MPI; 98-33331/29.
DR P-PSDB; W59783.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Pages 53-54; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 544 Bp; 166 A; 118 C; 112 G; 148 T;

Query Match 32.7%; Score 512; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 4.4e-249;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCACTCCGTTATCGCATTCATTTGGAACTATTGCAACTGCTA 127
DB 15 ATATGAAAAAATCACTCCGTTATCGCATTCATTTGGAACTATTGCAACTGCTA 74
QY 128 ATGGGGAGATTAACTGCAAGCAGCACTGCAACGGCACTCTGTGAACGCCGCCA 187
DB 75 ATGGGGAGATTAACTGCAAGCAGCACTGCAACGGCACTCTGTGAACGCCGCCA 134
QY 188 TCACCTTACATATAAGAGAGCGCTCCATTAATTAATGAGCAATGAGAACTGATA 247
DB 135 TCACCTTACATATAAGAGAGCGCTCCATTAATTAATGAGCAATGAGAACTGATA 194
QY 248 CAGAAATTAATGTTGGTGAAGCTTACTTGTGGCGCTATAAAGCAGAACACTAGCAT 307
DB 195 CAGAAATTAATGTTGGTGAAGCTTACTTGTGGCGCTATAAAGCAGAACACTAGCAT 254
QY 308 CTGTTAATTAAGATGCGCGGGGTATGCCATGTAATTAATTAATTAATTAATTA 367
DB 255 CTGTTAATTAAGATGCGCGGGGTATGCCATGTAATTAATTAATTAATTAATTA 314
QY 368 GAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 427
DB 315 GAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 374
QY 428 CTCCTAAGTAAACGGTGAAGACCTTGTGGGGATGACGTCTTGTGGCTACGGGACGCC 487

Db 375 CTCTTAAGGTAAAGGAGAGACCTTGGGGAGTGAAGTGGTCTTGCTAGGGGACACC 434
 488 AGATTTCTTTGTCGGTCATGTTTCCAAAGGGGTAACCTGACAGTAATACA 547
 Db 435 AGGATTTCTTTGTCGGTCATGTTTCCAAAGGGGTAACCTGACAGTAATACA 494
 QY 548 CTGATGCTGTAAACCGTAACCGTATCTAACCA 579
 Db 495 CTGATGCTGTAAACCGTAACCGTATCTAACCA 526

RESULT 11

VA1595 standard; DNA; 510 BP.
 ID VA1595;
 AC 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nYpF1(b)sec544.
 KW F1 antigen; plasmid; vaccine; plague; ds.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HEK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Page 53; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 510 BP; 153 A; 114 C; 106 G; 137 T;

Query Match 32.6%; Score 510; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4.5e-248;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 ATGAAAAAATCACTCCGTTATCGCATTCATTTATGGACATTTGCAATGCTAT 129
 Db 1 ATGAAAAAATCACTCCGTTATCGCATTCATTTATGGACATTTGCAATGCTAT 60
 QY 130 GCGGACAGTTTAACTGCAAGACCACTGCAAGCGCACTCTGTTGAACCAAGCCGCATC 189
 Db 61 GCGGACAGTTTAACTGCAAGACCACTGCAAGCGCACTCTGTTGAACCAAGCCGCATC 120
 QY 190 ACTTTACATTAAGGAAGGGCTCCATTTACATTAAGGAAGGAATGATACA 249
 Db 121 ACTTTACATTAAGGAAGGGCTCCATTTACATTAAGGAAGGAATGATACA 180
 QY 250 GAATTAATGTTGGTACGCTTACTCTTGCGGCTATAAAGAGNACACTAGCATCT 309
 Db 181 GAATTAATGTTGGTACGCTTACTCTTGCGGCTATAAAGAGNACACTAGCATCT 240
 QY 310 GTTAACCTTACAGATGCCGGGGTGATCCATGTACTTAACATTTACTCTCAGATGA 369
 Db 241 GTTAACCTTACAGATGCCGGGGTGATCCATGTACTTAACATTTACTCTCAGATGA 300
 QY 370 AATAACCAACCAATTCACACAAAGTGTATGGCAAGATTTAGAGATTTGATCTCT 429
 Db 301 AATAACCAACCAATTCACACAAAGTGTATGGCAAGATTTAGAGATTTGATCTCT 360
 QY 430 CCTAAGGTAAAGGCTGAGAACCTTGTGGGGGATGACGTCGCTTGGCTAGCGGACCCAG 489
 Db 361 CCTAAGGTAAAGGCTGAGAACCTTGTGGGGGATGACGTCGCTTGGCTAGCGGACCCAG 420
 QY 490 GATTTCTTTGTCGGTCATGTTTCCAAAGGGGTAACCTGACAGTAATACA 549
 Db 421 GATTTCTTTGTCGGTCATGTTTCCAAAGGGGTAACCTGACAGTAATACA 480

QY 550 GATGCTGTAAACCGTAACCGTATCTAACCA 579
 Db 481 GATGCTGTAAACCGTAACCGTATCTAACCA 510

RESULT 12

VA1601 standard; DNA; 450 BP.
 ID VA1601;
 AC 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nYpF1mat450.
 KW F1 antigen; plasmid; vaccine; plague; ds.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HEK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Page 61; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 450 BP; 134 A; 102 C; 96 G; 118 T;

Query Match 28.6%; Score 448; DB 1; Length 450;
 Best Local Similarity 100.0%; Pred. No. 9.3e-217;

Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 GCGAGATTTAATCTGCAAGACCACTGCAAGCGCACTCTGTTGAACAGCCGCATCAC 191
 Db 3 GCGAGATTTAATCTGCAAGACCACTGCAAGCGCACTCTGTTGAACAGCCGCATCAC 62
 QY 192 TCTTACATTAAGGAAGGGCTCCATTTACATTAAGGAAGGAATGATACA 251
 Db 63 TCTTACATTAAGGAAGGGCTCCATTTACATTAAGGAAGGAATGATACA 122
 QY 252 ATTACTTGTGTGACGCTTACTCTTGCGGCTATAAAGAGNACACTAGCATCTGT 311
 Db 123 ATTACTTGTGTGACGCTTACTCTTGCGGCTATAAAGAGNACACTAGCATCTGT 182
 QY 312 TAACTTTACAGATGCCGGGGTGATCCATGTACTTAACATTTACTCTCAGATGA 371
 Db 183 TAACTTTACAGATGCCGGGGTGATCCATGTACTTAACATTTACTCTCAGATGA 242
 QY 372 TAACACCAATTCACACAAAGTGTATGGCAAGATTTAGAGATTTGATCTCTCC 431
 Db 243 TAACACCAATTCACACAAAGTGTATGGCAAGATTTAGAGATTTGATCTCTCC 302
 QY 432 TAAAGTAAAGGCTGAGAACCTTGTGGGGGATGACGTCGCTTGGCTAGCGGACCCAGA 491
 Db 303 TAAAGTAAAGGCTGAGAACCTTGTGGGGGATGACGTCGCTTGGCTAGCGGACCCAGA 362
 QY 492 TTTGTTTGTGTCGTCATGTTTCCAAAGGGGTAACCTGACAGTAATACA 551
 Db 363 TTTGTTTGTGTCGTCATGTTTCCAAAGGGGTAACCTGACAGTAATACA 422
 QY 552 TGCTGTAAACCGTAACCGTATCTAACCA 579
 Db 423 TGCTGTAAACCGTAACCGTATCTAACCA 450

RESULT 13

VA1600 standard; DNA; 474 BP.
 ID VA1600;
 AC VA1600;

DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nypfimat474.
 KW F1 antigen; plasmid; vaccine; plaque; ds.
 OS Yersinia pestis.
 FT Key Location/Qualifiers
 FT CDS 7,459
 FT /tag= a
 FT /product= "F1 antigen"
 WO9824912-A2.
 PN 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR MPI: 98-333331/29.
 DR P-PSDB: W59787.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Pages 59-60; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 474 BP; 143 A; 106 C; 99 G; 126 T;

Query Match 28.6%; Score 448; DB 1; Length 474;
 Best Local Similarity 100.0%; Pred. No. 9.3e-217; Indels 0; Gaps 0;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 GGCAGATTTAACTGACAGACACACGCAAGGCAACTCTTTGTAACCAAGCCGCAATCAC 191
 DB 9 GGCAGATTTAACTGACAGACACACGCAAGGCAACTCTTTGTAACCAAGCCGCAATCAC 68
 QY 192 TCTTAATATTAAGAGAGCGCCCTCAATTAATATGAGAAATGAAACATCGATACAGA 251
 DB 69 TCTTAATATTAAGAGAGCGCCCTCAATTAATATGAGAAATGAAACATCGATACAGA 128
 QY 252 AATTACTGTTGAGACCTTACTCTTGGCGGCTATTAACAGAGAACCACTAGACATCTGT 311
 DB 129 AATTACTGTTGAGACCTTACTCTTGGCGGCTATTAACAGAGAACCACTAGACATCTGT 188
 QY 312 TAACTTTACAGATGCGCGGGGTGATCCCATGTAATTAATTAATCTTCTCAGATGGAAA 371
 DB 189 TAACTTTACAGATGCGCGGGGTGATCCCATGTAATTAATTAATCTTCTCAGATGGAAA 248
 QY 372 TAACCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 431
 DB 249 TAACCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 308
 QY 432 TAAGTAAGAGGTGAGAACCTTGTGGGGATGACGTCGTTGGCTACAGGGACAGCAAGA 491
 DB 309 TAAGTAAGAGGTGAGAACCTTGTGGGGATGACGTCGTTGGCTACAGGGACAGCAAGA 368
 QY 492 TTTCTTTGTTGCTCAATTTGGTTCCAAAGCGGTAAACTTGACAGAGGTAAATACACTGA 551
 DB 369 TTTCTTTGTTGCTCAATTTGGTTCCAAAGCGGTAAACTTGACAGAGGTAAATACACTGA 428
 QY 552 TGCTGTACCGGTACCGGTATCTAACCA 579
 DB 429 TGCTGTACCGGTACCGGTATCTAACCA 456

RESULT 14
 ID V41609 standard; DNA; 447 BP.
 AC V41609;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nypfimat447.
 KW F1 antigen; plasmid; vaccine; plaque; ds.
 OS Yersinia pestis.
 FT Key Location/Qualifiers
 FT CDS 1,576
 FT /tag= a
 FT /product= "F1 antigen"
 WO9824912-A2.
 PN 11-JUN-1998.

FT CDS 1,447
 FT /tag= a
 FT /product= "F1 antigen"
 WO9824912-A2.
 PN 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR MPI: 98-333331/29.
 DR P-PSDB: W59788.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Page 64; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 447 BP; 133 A; 102 C; 95 G; 117 T;

Query Match 28.5%; Score 447; DB 1; Length 447;
 Best Local Similarity 100.0%; Pred. No. 3e-216; Indels 0; Gaps 0;
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 GCAGATTTAACTGACAGACACACGCAAGGCAACTCTTTGTAACCAAGCCGCAATCAC 192
 DB 1 GCAGATTTAACTGACAGACACACGCAAGGCAACTCTTTGTAACCAAGCCGCAATCAC 60
 QY 193 CTTCATATTAAGAGAGCGCCCTCAATTAATATGAGAAATGAAACATCGATACAGA 252
 DB 61 CTTCATATTAAGAGAGCGCCCTCAATTAATATGAGAAATGAAACATCGATACAGA 120
 QY 253 TTAATGTTGTTGCTTACTCTTGGCGGCTATTAACAGAGAACCACTAGACATCTGT 312
 DB 121 TTAATGTTGTTGCTTACTCTTGGCGGCTATTAACAGAGAACCACTAGACATCTGT 180
 QY 313 AACTTTACAGATGCGCGGGGTGATCCCATGTAATTAATTAATTAATTAATTAATTAAT 372
 DB 181 AACTTTACAGATGCGCGGGGTGATCCCATGTAATTAATTAATTAATTAATTAATTAAT 240
 QY 373 AACCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 432
 DB 241 AACCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
 QY 433 AAGTAAGAGGTGAGAACCTTGTGGGGATGACGTCGTTGGCTACAGGGACAGCAAGAT 492
 DB 301 AAGTAAGAGGTGAGAACCTTGTGGGGATGACGTCGTTGGCTACAGGGACAGCAAGAT 360
 QY 493 TTTCTTTGTTGCTCAATTTGGTTCCAAAGCGGTAAACTTGACAGAGGTAAATACACTGAT 552
 DB 361 TTTCTTTGTTGCTCAATTTGGTTCCAAAGCGGTAAACTTGACAGAGGTAAATACACTGAT 420
 QY 553 GCTGTACCGGTACCGGTATCTAACCA 579
 DB 421 GCTGTACCGGTACCGGTATCTAACCA 447

RESULT 15
 ID V41598 standard; DNA; 576 BP.
 AC V41598;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nypfimat576.
 KW F1 antigen; plasmid; vaccine; plaque; ds.
 OS Yersinia pestis.
 FT Key Location/Qualifiers
 FT CDS 1,576
 FT /tag= a
 FT /product= "F1 antigen"
 WO9824912-A2.
 PN 11-JUN-1998.

PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI; 98-333331/29.
DR P-PSDB; W59785.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Pages 55-56; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 576 BP; 179 A; 107 C; 104 G; 186 T;

Query Match 25.7%; Score 402; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.6e-193;

Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 ATGAAAAATCACTCCGTTATGCCATTCATTTTGGAACTATGCAACTGCTAAT 129
DB 1 ATGAAAAATCACTCCGTTATGCCATTCATTTTGGAACTATGCAACTGCTAAT 60
QY 130 GCGGCAATTTACTGCAAGACCACTGCAAGGCACTCTTGTGAACCGCCGATC 189
DB 61 GCGGCAATTTACTGCAAGACCACTGCAAGGCACTCTTGTGAACCGCCGATC 120
QY 190 ACTCTTACATATAAGAGGCGCTCAATTACAATTATGACAATGGAACATGATACA 249
DB 121 ACTCTTACATATAAGAGGCGCTCAATTACAATTATGACAATGGAACATGATACA 180
QY 250 GAATTAATTTGTTGTAAGCTTACTCTTGGCGGCTATATAAGAGACCACTAGCATCT 309
DB 181 GAATTAATTTGTTGTAAGCTTACTCTTGGCGGCTATATAAGAGACCACTAGCATCT 240
QY 310 GTTAACTTTACAGATGCGGCGGCTGATCCCATGTACTTAACATTTCTCAGGATGA 369
DB 241 GTTAACTTTACAGATGCGGCGGCTGATCCCATGTACTTAACATTTCTCAGGATGA 300
QY 370 AATTAACCAATTCACCTACAAAAGTATGGAAGATTTCTAGAGATTTGATATCTCT 429
DB 301 AATTAACCAATTCACCTACAAAAGTATGGAAGATTTCTAGAGATTTGATATCTCT 360
QY 430 CCTAAGGTAAACGCTGAGAACCTTGTGGGGATGACGTCGTC 471
DB 361 CCTAAGGTAAACGCTGAGAACCTTGTGGGGATGACGTCGTC 402

Search completed: August 22, 2000, 16:30:24
Job time: 3076 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 15:36:43 ; Search time 46.98 Seconds
(Without alignments)

4585.194 Million cell updates/sec

Title: US-08-699-716a-1

Perfect score: 1566
Sequence: 1 ATGGGCGCATCATCATCATCA.....ATGCACGCTGCTGAATCA 1566

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 243080 seqs, 68777915 residues

Word size: 0

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database:

Issued_Patents_NA:*
1: /cgn2-6/ptodata/1/1na/5A.COMB.seq:*
2: /cgn2-6/ptodata/1/1na/5B.COMB.seq:*
3: /cgn2-6/ptodata/1/1na/5C.COMB.seq:*
4: /cgn2-6/ptodata/1/1na/5D.COMB.seq:*
5: /cgn2-6/ptodata/1/1na/6.COMB.seq:*
6: /cgn2-6/ptodata/1/1na/PTUS.COMB.seq:*
7: /cgn2-6/ptodata/1/1na/Backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	927	59.2	1014	4 US-08-913-477-1	Sequence 1, Appl
2	927	59.2	1462	4 US-08-913-477-16	Sequence 16, Appl
3	927	59.2	1530	4 US-08-913-477-22	Sequence 22, Appl
4	923	58.9	1014	4 US-08-913-477-3	Sequence 3, Appl
5	411	26.2	547	4 US-08-913-477-20	Sequence 20, Appl
6	398	25.4	541	4 US-08-913-477-10	Sequence 10, Appl
7	68	4.3	375	5 US-08-870-370-8	Sequence 8, Appl
8	68	4.3	375	5 US-08-870-370-9	Sequence 9, Appl
9	68	4.3	375	5 US-08-870-370-8	Sequence 8, Appl
10	53	3.4	1402	2 US-08-480-604A-25	Sequence 25, Appl
11	53	3.4	1402	4 US-08-405-486A-25	Sequence 25, Appl
12	35	2.2	47	4 US-08-600-783-15	Sequence 25, Appl
13	30	1.9	8920	3 US-08-446-855A-11	Sequence 11, Appl
14	29	1.9	4084	5 US-08-866-340-1	Sequence 1, Appl
15	28	1.8	39	3 US-08-432-871C-44	Sequence 44, Appl
16	28	1.8	88	6 PCT-US95-10973A-36	Sequence 36, Appl
17	28	1.8	287	4 US-08-347-563A-11	Sequence 11, Appl
18	28	1.8	287	4 US-08-292-345B-11	Sequence 11, Appl
19	28	1.8	287	4 US-08-485-942A-11	Sequence 11, Appl
20	28	1.8	732	6 PCT-US96-05611A-29	Sequence 29, Appl
21	28	1.8	2136	5 US-08-946-475-8	Sequence 8, Appl
22	28	1.8	6693	4 US-08-147-777-2	Sequence 2, Appl
23	28	1.8	6693	5 US-08-452-872-2	Sequence 2, Appl
24	28	1.8	6693	6 PCT-US93-03985-2	Sequence 2, Appl
25	28	1.8	10747	4 US-08-147-777-1	Sequence 1, Appl
26	28	1.8	10747	5 US-08-452-872-1	Sequence 1, Appl

27	28	1.8	10747	6 PCT-US93-03985-1	Sequence 1, Appl
28	28	1.8	24979	4 US-08-147-777-3	Sequence 3, Appl
29	28	1.8	24979	5 US-08-452-872-3	Sequence 3, Appl
30	28	1.8	24979	6 PCT-US93-03985-3	Sequence 3, Appl
31	25	1.6	225	4 US-08-913-477-9	Sequence 9, Appl
32	24	1.5	1667	2 US-08-485-284A-1	Sequence 1, Appl
33	23	1.5	32	4 US-08-913-477-7	Sequence 7, Appl
34	23	1.5	797	4 US-08-752-132-1	Sequence 1, Appl
35	23	1.5	5434	3 US-08-841-349-1	Sequence 1, Appl
36	23	1.5	7635	2 US-08-455-533A-23	Sequence 1, Appl
37	23	1.5	7635	3 US-08-455-533A-23	Sequence 23, Appl
38	23	1.5	7635	3 US-08-193-078B-1	Sequence 1, Appl
39	23	1.5	7635	3 US-08-193-078B-29	Sequence 29, Appl
40	23	1.5	7635	3 US-08-223-305C-1	Sequence 1, Appl
41	23	1.5	7635	3 US-08-223-305C-23	Sequence 23, Appl
42	23	1.5	7635	3 US-08-129-097D-1	Sequence 1, Appl
43	22	1.4	2625	3 US-08-759-945-1	Sequence 1, Appl
44	22	1.4	7125	3 US-07-745-206A-1	Sequence 1, Appl
45	22	1.4	7125	3 US-08-311-363-1	Sequence 1, Appl
46	21	1.3	21	4 US-08-863-639A-44	Sequence 44, Appl
47	21	1.3	21	4 US-08-863-639A-45	Sequence 45, Appl
48	21	1.3	21	4 US-08-863-639A-49	Sequence 49, Appl
49	21	1.3	21	4 US-08-863-639A-65	Sequence 65, Appl
50	21	1.3	21	4 US-08-863-639A-82	Sequence 82, Appl
51	21	1.3	21	4 US-08-863-639A-86	Sequence 86, Appl
52	21	1.3	47	3 US-08-762-433-5	Sequence 5, Appl
53	21	1.3	47	3 US-09-001-219-5	Sequence 5, Appl
54	21	1.3	1475	2 US-08-097-938-1	Sequence 1, Appl
55	21	1.3	1475	2 US-08-476-000-1	Sequence 1, Appl
56	21	1.3	1475	2 US-08-472-840-1	Sequence 1, Appl
57	21	1.3	1475	2 US-08-472-840-1	Sequence 1, Appl
58	21	1.3	1475	2 US-08-476-976-1	Sequence 1, Appl
59	21	1.3	1497	5 US-08-474-410-1	Sequence 5, Appl
60	21	1.3	1497	6 PCT-US96-06511-5	Sequence 5, Appl
61	21	1.3	2732	2 US-08-476-000-50	Sequence 60, Appl
62	21	1.3	2732	2 US-08-472-840-50	Sequence 60, Appl
63	21	1.3	2732	2 US-08-476-976-60	Sequence 60, Appl
64	21	1.3	2732	5 US-08-474-410-60	Sequence 60, Appl
65	20	1.3	28	4 US-08-913-477-19	Sequence 19, Appl
66	20	1.3	40	7 US-07-743-245-1	Sequence 1, Appl
67	20	1.3	1297	7 5427925-21	Patent No. 5187077
68	20	1.3	1297	7 5427925-21	Patent No. 5187077
69	20	1.3	2469	1 US-07-997-133-2	Sequence 2, Appl
70	20	1.3	2469	1 US-08-459-296-1	Sequence 1, Appl
71	20	1.3	2469	6 US-07-997-133-2	Sequence 2, Appl
72	20	1.3	2662	3 US-08-451-822A-14	Sequence 14, Appl
73	20	1.3	2733	6 PCT-US96-00331-14	Sequence 14, Appl
74	20	1.3	5194	3 US-08-642-846-1	Sequence 1, Appl
75	20	1.3	30	2 US-08-737-757-7	Sequence 7, Appl
76	19	1.2	30	2 US-08-737-757-8	Sequence 8, Appl
77	19	1.2	33	2 US-08-482-882-29	Sequence 29, Appl
78	19	1.2	33	2 US-08-483-389-29	Sequence 29, Appl
79	19	1.2	33	2 US-08-483-389-29	Sequence 29, Appl
80	19	1.2	33	3 US-08-487-113D-29	Sequence 29, Appl
81	19	1.2	33	3 US-08-473-503-29	Sequence 29, Appl
82	19	1.2	33	3 US-08-483-922-29	Sequence 29, Appl
83	19	1.2	33	4 US-08-720-400A-29	Sequence 29, Appl
84	19	1.2	33	4 US-08-720-400A-30	Sequence 30, Appl
85	19	1.2	40	2 US-08-482-882-30	Sequence 30, Appl
86	19	1.2	40	2 US-08-483-389-30	Sequence 30, Appl
87	19	1.2	40	3 US-08-487-113D-30	Sequence 30, Appl
88	19	1.2	40	3 US-08-473-503-30	Sequence 30, Appl
89	19	1.2	40	3 US-08-483-922-30	Sequence 30, Appl
90	19	1.2	40	4 US-08-483-922-30	Sequence 30, Appl
91	19	1.2	40	4 US-08-720-400A-30	Sequence 30, Appl
92	19	1.2	41	2 US-08-714-017-30	Sequence 30, Appl
93	19	1.2	45	5 US-08-525-654A-139	Sequence 139, App
94	19	1.2	49	4 US-08-481-435-19	Sequence 19, Appl
95	19	1.2	49	4 US-09-061-337-19	Sequence 19, Appl
96	19	1.2	49	4 US-09-122-139-19	Sequence 19, Appl
97	19	1.2	51	4 US-09-340-991-19	Sequence 19, Appl
98	19	1.2	51	4 US-09-061-337-13	Sequence 13, Appl
99	19	1.2	51	5 US-09-122-129-13	Sequence 13, Appl
				US-09-340-991-13	Sequence 13, Appl

100	19	1.2	53	1	US-08-049-264C-12	Sequence 12, Appl	173	18	1.1	3661	1	US-08-105-483-221	Sequence 221, App
101	19	1.2	53	1	US-08-476-562-12	Sequence 12, Appl	174	18	1.1	3661	1	US-08-303-124-7	Sequence 7, Appl1
102	19	1.2	53	2	US-08-479-723A-12	Sequence 12, Appl	175	18	1.1	3661	1	US-08-204-729-7	Sequence 7, Appl1
103	19	1.2	53	6	PCT-US94-04310-12	Sequence 12, Appl	176	18	1.1	3661	2	US-08-709-209-221	Sequence 221, App
104	19	1.2	57	4	US-08-662-227-39	Sequence 39, Appl	177	18	1.1	3661	2	US-08-458-101-221	Sequence 221, App
105	19	1.2	96	5	US-08-737-336-2	Sequence 23, Appl1	178	18	1.1	3661	2	US-08-480-697B-7	Sequence 7, Appl1
106	19	1.2	105	4	US-08-687-865A-23	Sequence 23, Appl1	179	18	1.1	4453	3	US-08-770-761A-6	Sequence 6, Appl1
107	19	1.2	129	4	US-08-579-865-38	Sequence 38, Appl	180	18	1.1	4540	2	US-08-770-761A-6	Sequence 6, Appl1
108	19	1.2	129	4	US-08-680-876-38	Sequence 38, Appl	181	18	1.1	6387	1	US-07-721-775A-1	Sequence 1, Appl1
109	19	1.2	153	4	US-08-679-865-41	Sequence 41, Appl	182	18	1.1	6387	1	US-08-339-658-1	Sequence 1, Appl1
110	19	1.2	153	4	US-08-680-876-41	Sequence 41, Appl	183	17	1.1	5463174-1	7	5463174-1	Patent No. 5463174
111	19	1.2	234	4	US-08-687-865A-20	Sequence 20, Appl	184	17	1.1	27	7	5463174-2	Patent No. 5463174
112	19	1.2	466	1	US-08-457-245-20	Sequence 20, Appl	185	17	1.1	41	3	US-08-632-877C-43	Sequence 43, Appl
113	19	1.2	927	1	US-08-99-568-1	Sequence 1, Appl1	186	17	1.1	303	5	US-08-621-018B-5	Sequence 5, Appl1
114	19	1.2	927	2	US-08-793-958-1	Sequence 1, Appl1	187	17	1.1	1138	4	US-08-993-225-5	Sequence 5, Appl1
115	19	1.2	1056	1	US-08-701-191A-5	Sequence 5, Appl1	188	17	1.1	1183	4	US-08-188-582-25	Sequence 25, Appl
116	19	1.2	1240	4	US-09-061-337-9	Sequence 9, Appl1	189	17	1.1	1183	1	US-08-646-715-25	Sequence 25, Appl
117	19	1.2	1240	4	US-09-122-129-9	Sequence 9, Appl1	190	17	1.1	1282	1	US-08-211-942-16	Sequence 16, Appl
118	19	1.2	1240	5	US-09-340-991-9	Sequence 9, Appl1	191	17	1.1	1328	1	US-08-290-448A-58	Sequence 58, Appl
119	19	1.2	1240	5	US-08-737-336-5	Sequence 5, Appl1	192	17	1.1	1328	2	US-08-175-069A-58	Sequence 58, Appl
120	19	1.2	1599	5	US-08-737-336-5	Sequence 5, Appl1	193	17	1.1	1328	2	US-08-290-448A-58	Sequence 58, Appl
121	19	1.2	1740	4	US-08-335-760-3	Sequence 3, Appl1	194	17	1.1	1349	2	US-08-290-448A-73	Sequence 73, Appl
122	19	1.2	1896	4	US-08-605-541B-11	Sequence 11, Appl	195	17	1.1	1349	2	US-08-290-448A-73	Sequence 73, Appl
123	19	1.2	2081	4	US-09-096-982-7	Sequence 7, Appl1	196	17	1.1	1349	2	US-08-175-069A-73	Sequence 73, Appl
124	19	1.2	2081	6	US-08-653-650A-7	Sequence 7, Appl1	197	17	1.1	1349	2	US-08-175-069A-73	Sequence 73, Appl
125	19	1.2	3090	6	PCT-US93-06251-7	Sequence 7, Appl1	198	17	1.1	1349	2	US-08-175-069A-73	Sequence 73, Appl
126	19	1.2	3255	2	US-08-717-515-5	Sequence 5, Appl1	199	17	1.1	1349	2	US-08-175-069A-73	Sequence 73, Appl
127	19	1.2	3704	4	US-09-014-969-20	Sequence 20, Appl	200	17	1.1	2188	1	US-09-221-235-4	Sequence 4, Appl1
128	19	1.2	3831	2	US-08-717-515-7	Sequence 7, Appl1	201	17	1.1	2188	1	US-08-665-662F-10	Sequence 10, Appl
129	19	1.2	6934	4	US-08-015-973-2	Sequence 2, Appl1	202	17	1.1	2295	5	US-08-375-300-3	Sequence 3, Appl1
130	19	1.2	6934	4	US-08-448-164-2	Sequence 2, Appl1	203	17	1.1	2295	5	US-09-177-431-3	Sequence 3, Appl1
131	19	1.2	7286	6	PCT-US95-11684-1	Sequence 1, Appl1	204	17	1.1	2295	5	PCT-US95-16930-3	Sequence 3, Appl1
132	18	1.1	18	5	US-08-047-837-21	Sequence 21, Appl	205	17	1.1	3074	6	PCT-US93-11725-3	Sequence 3, Appl1
133	18	1.1	18	5	US-09-031-897-10	Sequence 10, Appl	206	17	1.1	3074	6	US-08-118-101A-1	Sequence 1, Appl1
134	18	1.1	35	1	US-08-049-264C-10	Sequence 10, Appl	207	17	1.1	4080	5	US-08-375-300-1	Sequence 1, Appl1
135	18	1.1	35	1	US-08-476-562-10	Sequence 10, Appl	208	17	1.1	4080	5	US-09-177-431-1	Sequence 1, Appl1
136	18	1.1	35	6	US-08-479-723A-10	Sequence 10, Appl	209	17	1.1	5183	3	PCT-US95-16930-1	Sequence 1, Appl1
137	18	1.1	40	3	PCT-US94-04310-10	Sequence 10, Appl	210	17	1.1	5183	3	US-08-459-568-3	Sequence 3, Appl1
138	18	1.1	43	5	US-09-136-421-1	Sequence 5, Appl1	211	17	1.1	5183	3	US-08-399-411-3	Sequence 3, Appl1
139	18	1.1	45	5	US-08-481-435-40	Sequence 40, Appl	212	17	1.1	5183	3	US-08-516-859A-3	Sequence 3, Appl1
140	18	1.1	49	2	US-08-052-157-1	Sequence 2, Appl1	213	17	1.1	9468	2	US-08-325-547-9	Sequence 2, Appl1
141	18	1.1	51	1	US-08-442-334A-4	Sequence 4, Appl1	214	17	1.1	9472	2	US-08-325-547-9	Sequence 2, Appl1
142	18	1.1	51	1	US-08-444-581B-4	Sequence 4, Appl1	215	17	1.1	13875	3	US-08-734-344-1	Sequence 1, Appl1
143	18	1.1	51	1	US-08-446-088A-4	Sequence 4, Appl1	216	16	1.0	21	5	US-08-851-350-32	Sequence 32, Appl
144	18	1.1	51	3	US-08-642-541-7	Sequence 7, Appl1	217	16	1.0	21	5	US-08-851-350-32	Sequence 32, Appl
145	18	1.1	53	4	US-08-687-865A-10	Sequence 10, Appl	218	16	1.0	24	4	US-08-600-783-16	Sequence 16, Appl
146	18	1.1	55	4	US-08-687-865A-11	Sequence 11, Appl	219	16	1.0	29	1	US-08-276-852-140	Sequence 140, App
147	18	1.1	58	3	US-08-484-993B-39	Sequence 39, Appl	220	16	1.0	29	1	US-08-276-852-140	Sequence 140, App
148	18	1.1	58	4	US-08-484-158B-39	Sequence 39, Appl	221	16	1.0	29	2	US-08-899-575-139	Sequence 139, App
149	18	1.1	58	4	US-08-484-596A-39	Sequence 39, Appl	222	16	1.0	29	2	US-08-899-575-140	Sequence 140, App
150	18	1.1	58	4	US-08-480-150A-39	Sequence 39, Appl	223	16	1.0	29	2	US-08-899-575-139	Sequence 139, App
151	18	1.1	58	5	US-08-458-731-39	Sequence 39, Appl	224	16	1.0	29	2	US-08-899-575-140	Sequence 140, App
152	18	1.1	58	5	US-08-149-223A-39	Sequence 39, Appl	225	16	1.0	29	6	PCT-US95-08743-139	Sequence 139, App
153	18	1.1	688	2	US-08-139-937-9	Sequence 9, Appl1	226	16	1.0	29	6	PCT-US95-08743-139	Sequence 139, App
154	18	1.1	688	6	PCT-US93-11310-9	Sequence 9, Appl1	227	16	1.0	39	5	US-08-792-832A-8	Sequence 8, Appl
155	18	1.1	1293	3	US-08-722-184-7	Sequence 7, Appl1	228	16	1.0	42	5	US-09-136-421-2	Sequence 2, Appl1
156	18	1.1	2109	3	US-08-617-801A-5	Sequence 5, Appl1	229	16	1.0	42	5	US-09-136-421-2	Sequence 2, Appl1
157	18	1.1	2220	3	US-08-617-801A-1	Sequence 1, Appl1	230	16	1.0	141	5	US-09-003-708A-2	Sequence 2, Appl1
158	18	1.1	2722	4	US-08-500-857A-7	Sequence 7, Appl1	231	16	1.0	220	1	US-08-276-852-112	Sequence 11, Appl
159	18	1.1	2809	1	US-08-484-105-3	Sequence 3, Appl1	232	16	1.0	220	1	US-08-276-852-112	Sequence 11, Appl
160	18	1.1	2809	1	US-08-484-106-3	Sequence 3, Appl1	233	16	1.0	220	1	US-08-276-852-112	Sequence 11, Appl
161	18	1.1	3635	3	US-08-553-436A-5	Sequence 5, Appl1	234	16	1.0	220	2	US-08-899-575-11	Sequence 11, Appl
162	18	1.1	3635	3	US-08-220-151-72	Sequence 72, Appl	235	16	1.0	220	2	US-08-899-575-112	Sequence 12, Appl
163	18	1.1	3635	1	US-08-303-124-12	Sequence 12, Appl	236	16	1.0	220	2	US-08-899-575-112	Sequence 12, Appl
164	18	1.1	3635	1	US-08-413-118-72	Sequence 72, Appl	237	16	1.0	220	6	PCT-US95-08743-11	Sequence 11, Appl
165	18	1.1	3635	2	US-08-324-657-48	Sequence 48, Appl	238	16	1.0	220	6	PCT-US95-08743-11	Sequence 11, Appl
166	18	1.1	3635	3	US-08-184-009-77	Sequence 77, Appl	239	16	1.0	315	4	US-08-487-031-21	Sequence 21, Appl
167	18	1.1	3635	3	US-08-417-210A-67	Sequence 67, Appl	240	16	1.0	526	2	US-08-700-575-1	Sequence 1, Appl1
168	18	1.1	3659	3	US-08-480-697B-12	Sequence 12, Appl	241	16	1.0	534	4	US-08-487-031-22	Sequence 22, Appl
169	18	1.1	3659	4	US-08-458-356-77	Sequence 77, Appl	242	16	1.0	536	4	US-08-487-031-19	Sequence 19, Appl
170	18	1.1	3659	5	US-08-473-446-72	Sequence 72, Appl	243	16	1.0	604	4	US-08-836-197-1	Sequence 1, Appl1
171	18	1.1	3660	2	US-08-475-063-25	Sequence 25, Appl	244	16	1.0	677	2	US-08-644-664B-3	Sequence 3, Appl1
172	18	1.1	3660	2	US-08-207-792-25	Sequence 25, Appl	245	16	1.0	677	4	US-08-761-277-3	Sequence 3, Appl1

C 246	16	1.0	909	4	US-08-487-031-18	Sequence 18, Appl	C 319	15	1.0	21	5	US-08-840-316-28	Sequence 28, Appl
C 247	16	1.0	1016	1	US-08-109-391A-3	Sequence 3, Appl1	C 320	15	1.0	21	6	PCT-US93-08849A-28	Sequence 28, Appl
C 248	16	1.0	1016	1	US-08-459-019A-3	Sequence 3, Appl1	C 321	15	1.0	21	6	PCT-US93-08849A-28	Sequence 28, Appl
C 249	16	1.0	1016	1	US-08-460-428A-3	Sequence 3, Appl1	C 322	15	1.0	30	1	US-08-255-0772-31	Sequence 31, Appl
C 250	16	1.0	1033	3	US-08-471-717-3	Sequence 3, Appl1	C 323	15	1.0	38	2	US-08-317-832A-32	Sequence 32, Appl
C 251	16	1.0	1158	2	US-08-891-254-4	Sequence 4, Appl1	C 324	15	1.0	38	2	US-08-317-832A-32	Sequence 32, Appl
C 252	16	1.0	1158	3	US-08-819-539-4	Sequence 4, Appl1	C 325	15	1.0	53	5	US-08-792-832A-17	Sequence 17, Appl
C 253	16	1.0	1158	6	PCT-US93-06243-4	Sequence 4, Appl1	C 326	15	1.0	57	5	US-08-634-060-30	Sequence 30, Appl
C 254	16	1.0	1158	6	PCT-US96-08819-4	Sequence 4, Appl1	C 327	15	1.0	81	1	US-07-939-501A-29	Sequence 29, Appl
C 255	16	1.0	1165	5	US-08-759-463-1	Sequence 1, Appl1	C 328	15	1.0	109	1	US-07-939-501A-27	Sequence 27, Appl
C 256	16	1.0	1165	5	US-07-876-380-8	Sequence 8, Appl1	C 329	15	1.0	176	4	US-08-687-080-82	Sequence 82, Appl
C 257	16	1.0	1185	1	US-08-049-783-5	Sequence 5, Appl1	C 330	15	1.0	247	2	US-08-691-641-4	Sequence 4, Appl1
C 258	16	1.0	1185	2	US-08-316-301A-9	Sequence 9, Appl1	C 331	15	1.0	305	1	US-08-207-904-13	Sequence 13, Appl
C 259	16	1.0	1185	3	US-08-904-278-9	Sequence 9, Appl1	C 332	15	1.0	342	2	US-08-486-013-56	Sequence 56, Appl
C 260	16	1.0	1189	1	US-07-781-034-4	Sequence 4, Appl1	C 333	15	1.0	342	3	US-08-482-279-56	Sequence 56, Appl
C 261	16	1.0	1189	6	PCT-US92-08328-4	Sequence 4, Appl1	C 334	15	1.0	342	3	US-08-342-268-56	Sequence 56, Appl
C 262	16	1.0	1203	3	US-08-739-943-2	Sequence 2, Appl1	C 335	15	1.0	342	5	US-09-015-568-56	Sequence 56, Appl
C 263	16	1.0	1203	3	US-08-200-724A-4	Sequence 4, Appl1	C 336	15	1.0	355	2	US-08-486-013-43	Sequence 43, Appl
C 264	16	1.0	1209	3	US-08-200-724A-4	Sequence 4, Appl1	C 337	15	1.0	355	3	US-08-482-279-43	Sequence 43, Appl
C 265	16	1.0	1227	6	PCT-US93-06243-3	Sequence 1, Appl1	C 338	15	1.0	355	3	US-08-342-268-43	Sequence 43, Appl
C 266	16	1.0	1287	3	US-08-683-007A-1	Sequence 3, Appl1	C 339	15	1.0	355	5	US-09-015-568-43	Sequence 43, Appl
C 267	16	1.0	1288	3	US-08-200-724A-3	Sequence 3, Appl1	C 340	15	1.0	466	5	US-08-928-799A-1	Sequence 1, Appl
C 268	16	1.0	1288	4	US-09-030-270A-4	Sequence 4, Appl1	C 341	15	1.0	466	5	US-08-486-013-62	Sequence 62, Appl
C 269	16	1.0	1320	1	US-08-250-448A-75	Sequence 75, Appl	C 342	15	1.0	419	2	US-08-482-279-62	Sequence 62, Appl
C 270	16	1.0	1320	2	US-08-290-448A-75	Sequence 75, Appl	C 343	15	1.0	419	3	US-08-342-268-62	Sequence 62, Appl
C 271	16	1.0	1320	2	US-08-175-069A-75	Sequence 75, Appl	C 344	15	1.0	419	5	US-09-015-568-62	Sequence 62, Appl
C 272	16	1.0	1400	3	US-08-001-078A-2	Sequence 2, Appl1	C 345	15	1.0	422	2	US-08-486-013-41	Sequence 41, Appl
C 273	16	1.0	1400	4	US-08-463-218-2	Sequence 2, Appl1	C 346	15	1.0	422	2	US-08-486-013-41	Sequence 41, Appl
C 274	16	1.0	1400	6	PCT-US94-00253-2	Sequence 2, Appl1	C 347	15	1.0	422	3	US-08-482-279-41	Sequence 41, Appl
C 275	16	1.0	1697	1	US-08-181-271A-104	Sequence 104, App	C 348	15	1.0	422	3	US-08-482-279-45	Sequence 45, Appl
C 276	16	1.0	1697	1	US-08-449-315-104	Sequence 104, App	C 349	15	1.0	422	3	US-08-482-279-45	Sequence 45, Appl
C 277	16	1.0	1697	1	US-08-444-803-104	Sequence 104, App	C 350	15	1.0	422	3	US-08-342-268-45	Sequence 45, Appl
C 278	16	1.0	1697	1	US-08-449-043-104	Sequence 104, App	C 351	15	1.0	422	5	US-09-015-568-45	Sequence 45, Appl
C 279	16	1.0	1697	2	US-08-455-265A-104	Sequence 104, App	C 352	15	1.0	422	5	US-09-015-568-45	Sequence 45, Appl
C 280	16	1.0	1697	2	US-08-455-265A-104	Sequence 104, App	C 353	15	1.0	422	5	US-09-015-568-45	Sequence 45, Appl
C 281	16	1.0	1697	2	US-08-455-265A-104	Sequence 104, App	C 354	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 282	16	1.0	1697	2	US-08-455-265A-104	Sequence 104, App	C 355	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 283	16	1.0	1697	2	US-08-455-265A-104	Sequence 104, App	C 356	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 284	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 357	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 285	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 358	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 286	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 359	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 287	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 360	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 288	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 361	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 289	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 362	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 290	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 363	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 291	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 364	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 292	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 365	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 293	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 366	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 294	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 367	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 295	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 368	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 296	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 369	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 297	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 370	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 298	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 371	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 299	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 372	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 300	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 373	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 301	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 374	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 302	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 375	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 303	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 376	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 304	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 377	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 305	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 378	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 306	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 379	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 307	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 380	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 308	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 381	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 309	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 382	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 310	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 383	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 311	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 384	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 312	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 385	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 313	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 386	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 314	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 387	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 315	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 388	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 316	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 389	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 317	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 390	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1
C 318	16	1.0	1697	3	US-08-455-265A-104	Sequence 104, App	C 391	15	1.0	533	1	PCT-US91-01750-1	Sequence 1, Appl1

392	15	1.0	1924	3	US-08-756-317-1	Sequence 1, Appl1	C 465	15	1.0	3372	3	US-08-539-005-48	Sequence 48, Appl1
C 393	15	1.0	2030	2	US-08-530-950-1	Sequence 1, Appl1	466	15	1.0	3383	5	US-08-884-324-12	Sequence 12, Appl1
394	15	1.0	2060	1	US-07-721-761A-31	Sequence 31, Appl1	467	15	1.0	3440	27	US-08-471-791-27	Sequence 27, Appl1
395	15	1.0	2060	1	US-07-978-687-31	Sequence 31, Appl1	468	15	1.0	3440	6	PCT-US91-01746-27	Sequence 27, Appl1
396	15	1.0	2060	6	PCT-US91-01750-2	Sequence 2, Appl1	469	15	1.0	3465	2	US-08-578-649-3	Sequence 2, Appl1
397	15	1.0	2060	6	PCT-US91-05801-31	Sequence 31, Appl1	C 470	15	1.0	3625	2	US-08-356-354-3	Sequence 3, Appl1
398	15	1.0	2081	6	PCT-US91-01750-3	Sequence 3, Appl1	C 471	15	1.0	3625	4	US-08-778-656-3	Sequence 3, Appl1
C 399	15	1.0	2158	4	US-07-688-352C-3	Sequence 3, Appl1	C 472	15	1.0	3629	5	US-08-837-653-6	Sequence 6, Appl1
C 400	15	1.0	2158	4	US-08-474-379C-3	Sequence 3, Appl1	C 473	15	1.0	3704	4	US-08-702-153-1	Sequence 1, Appl1
C 401	15	1.0	2158	4	US-09-146-249A-3	Sequence 3, Appl1	C 474	15	1.0	3740	2	US-08-356-354-1	Sequence 1, Appl1
C 402	15	1.0	2158	6	PCT-US91-02714-3	Sequence 3, Appl1	C 475	15	1.0	3740	4	US-08-778-656-1	Sequence 1, Appl1
C 403	15	1.0	2173	1	US-08-036-210-6	Sequence 6, Appl1	C 476	15	1.0	3870	1	US-08-138-641-1	Sequence 1, Appl1
C 404	15	1.0	2173	4	US-08-449-609-6	Sequence 6, Appl1	C 477	15	1.0	3870	1	US-08-138-133-1	Sequence 1, Appl1
C 405	15	1.0	2245	3	US-08-617-801A-3	Sequence 3, Appl1	C 478	15	1.0	3871	4	US-08-599-455B-3	Sequence 3, Appl1
C 406	15	1.0	2248	1	US-08-369-780-1	Sequence 1, Appl1	C 479	15	1.0	3893	1	US-08-138-641-3	Sequence 3, Appl1
C 407	15	1.0	2248	1	US-08-475-682-1	Sequence 1, Appl1	C 480	15	1.0	3893	1	US-08-138-133-3	Sequence 3, Appl1
C 408	15	1.0	2248	2	US-08-780-833-1	Sequence 1, Appl1	C 481	15	1.0	3921	4	US-08-567-378-3	Sequence 3, Appl1
C 409	15	1.0	2248	2	US-08-636-036-1	Sequence 1, Appl1	C 482	15	1.0	3921	4	US-08-587-680A-3	Sequence 3, Appl1
C 410	15	1.0	2248	5	US-08-918-509-1	Sequence 1, Appl1	C 483	15	1.0	3973	4	US-08-036-210-21	Sequence 2, Appl1
C 411	15	1.0	2248	5	US-09-108-262-1	Sequence 1, Appl1	C 484	15	1.0	4002	1	US-08-449-609-21	Sequence 1, Appl1
C 412	15	1.0	2309	1	US-08-036-210-10	Sequence 10, Appl1	C 485	15	1.0	4002	1	US-08-331-488A-1	Sequence 1, Appl1
C 413	15	1.0	2309	4	US-08-449-609-10	Sequence 10, Appl1	C 486	15	1.0	4291	3	US-08-417-210A-81	Sequence 3, Appl1
C 414	15	1.0	2343	4	US-09-018-760-3	Sequence 3, Appl1	C 487	15	1.0	4291	3	US-08-233-788A-39	Sequence 39, Appl1
C 415	15	1.0	2344	5	US-08-893-852A-2	Sequence 2, Appl1	C 488	15	1.0	4406	1	US-08-484-106-1	Sequence 1, Appl1
C 416	15	1.0	2369	3	US-08-523-742-1	Sequence 1, Appl1	C 489	15	1.0	4437	2	US-08-559-303B-72	Sequence 2, Appl1
C 417	15	1.0	2372	1	US-07-903-103-1	Sequence 1, Appl1	C 490	15	1.0	4454	4	US-08-738-172-2	Sequence 4, Appl1
C 418	15	1.0	2372	1	US-08-044-619A-1	Sequence 1, Appl1	C 491	15	1.0	4471	5	US-08-615-942A-1	Sequence 1, Appl1
C 419	15	1.0	2372	1	US-08-283-911-1	Sequence 1, Appl1	C 492	15	1.0	4818	5	US-08-817-926-27	Sequence 27, Appl1
C 420	15	1.0	2372	1	US-08-245-500A-2	Sequence 2, Appl1	C 493	15	1.0	4940	1	US-08-484-106-1	Sequence 1, Appl1
C 421	15	1.0	2372	1	US-08-390-546-2	Sequence 2, Appl1	C 494	15	1.0	4940	1	US-08-484-106-1	Sequence 1, Appl1
C 422	15	1.0	2372	1	US-08-390-479A-2	Sequence 2, Appl1	C 495	15	1.0	5057	3	US-08-365-486A-12	Sequence 12, Appl1
C 423	15	1.0	2372	2	US-08-557-393-2	Sequence 2, Appl1	C 496	15	1.0	5108	1	US-07-642-002-1	Sequence 1, Appl1
C 424	15	1.0	2372	2	US-08-390-516C-2	Sequence 2, Appl1	C 497	15	1.0	5285	4	US-08-609-049A-29	Sequence 29, Appl1
C 425	15	1.0	2372	2	US-08-390-517A-2	Sequence 2, Appl1	C 498	15	1.0	5893	2	US-08-592-126-54	Sequence 54, Appl1
C 426	15	1.0	2372	2	US-08-390-515A-2	Sequence 2, Appl1	C 499	15	1.0	5893	4	US-08-687-080-44	Sequence 44, Appl1
C 427	15	1.0	2372	5	US-08-801-718-2	Sequence 2, Appl1	C 500	15	1.0	5992	3	US-08-475-891A-3	Sequence 3, Appl1
C 428	15	1.0	2372	5	US-09-073-567-1	Sequence 1, Appl1	C 501	15	1.0	6063	1	US-08-195-744-4	Sequence 4, Appl1
C 429	15	1.0	2382	3	US-08-468-556-4	Sequence 4, Appl1	C 502	15	1.0	6063	3	US-08-788-279-4	Sequence 4, Appl1
C 430	15	1.0	2433	1	US-08-248-021A-1	Sequence 1, Appl1	C 503	15	1.0	6226	2	US-08-542-363-1	Sequence 1, Appl1
C 431	15	1.0	2440	1	US-08-160-861-2	Sequence 2, Appl1	C 504	15	1.0	6585	5	US-08-746-111-4	Sequence 4, Appl1
C 432	15	1.0	2442	2	US-08-542-363-3	Sequence 3, Appl1	C 505	15	1.0	6605	2	US-08-769-309A-4	Sequence 4, Appl1
C 433	15	1.0	2508	4	US-08-850-993-1	Sequence 1, Appl1	C 506	15	1.0	6914	2	US-08-920-812-22	Sequence 22, Appl1
C 434	15	1.0	2508	4	US-08-160-861-1	Sequence 1, Appl1	C 507	15	1.0	6914	2	US-08-920-827-22	Sequence 22, Appl1
C 435	15	1.0	2592	4	US-08-036-210-14	Sequence 14, Appl1	C 508	15	1.0	6914	2	US-08-921-177-22	Sequence 22, Appl1
C 436	15	1.0	2692	4	US-08-449-609-14	Sequence 14, Appl1	C 509	15	1.0	6914	2	US-08-362-577C-22	Sequence 22, Appl1
C 437	15	1.0	2791	4	US-09-016-000-8	Sequence 8, Appl1	C 510	15	1.0	6914	2	US-08-920-828-22	Sequence 22, Appl1
C 438	15	1.0	2877	3	US-08-693-697-35	Sequence 35, Appl1	C 511	15	1.0	7168	5	US-08-840-316-4	Sequence 5, Appl1
C 439	15	1.0	2880	3	US-08-693-697-32	Sequence 32, Appl1	C 512	15	1.0	7168	6	PCT-US93-08849A-4	Sequence 4, Appl1
C 440	15	1.0	2880	3	US-08-588-526-1	Sequence 1, Appl1	C 513	15	1.0	7451	2	US-08-684-672-23	Sequence 23, Appl1
C 441	15	1.0	2936	4	US-08-338-172-1	Sequence 1, Appl1	C 514	15	1.0	7970	2	US-08-135-511-31	Sequence 31, Appl1
C 442	15	1.0	2974	1	US-08-408-486-2	Sequence 2, Appl1	C 515	15	1.0	7970	2	US-08-187-453-31	Sequence 31, Appl1
C 443	15	1.0	2980	3	US-08-463-081B-13	Sequence 13, Appl1	C 516	15	1.0	7970	2	US-08-562-988A-7	Sequence 7, Appl1
C 444	15	1.0	2980	3	US-08-461-379A-13	Sequence 13, Appl1	C 517	15	1.0	10216	3	US-08-875-154-1	Sequence 1, Appl1
C 445	15	1.0	2980	3	US-08-462-390B-13	Sequence 13, Appl1	C 518	15	1.0	10216	3	US-08-884-324-13	Sequence 13, Appl1
C 446	15	1.0	2980	5	US-08-463-074B-13	Sequence 13, Appl1	C 519	15	1.0	11464	5	US-07-920-281C-1	Sequence 1, Appl1
C 447	15	1.0	2980	5	US-08-465-583C-13	Sequence 13, Appl1	C 520	15	1.0	11517	2	US-08-222-617A-3	Sequence 3, Appl1
C 448	15	1.0	2980	5	US-08-652-446-13	Sequence 13, Appl1	C 521	15	1.0	11601	3	US-08-222-617A-3	Sequence 3, Appl1
C 449	15	1.0	2991	2	US-08-355-888A-6	Sequence 6, Appl1	C 522	15	1.0	11601	3	US-08-222-617A-24	Sequence 24, Appl1
C 450	15	1.0	2991	3	US-08-588-190-1	Sequence 1, Appl1	C 523	15	1.0	13011	4	US-08-791-849A-14	Sequence 14, Appl1
C 451	15	1.0	2991	3	US-08-693-697-6	Sequence 6, Appl1	C 524	15	1.0	13999	5	US-08-444-644-24	Sequence 24, Appl1
C 452	15	1.0	3000	4	US-08-640-389A-1	Sequence 1, Appl1	C 525	15	1.0	19307	4	US-08-836-022A-10	Sequence 10, Appl1
C 453	15	1.0	3000	4	US-08-184-252A-1	Sequence 1, Appl1	C 526	15	1.0	22846	5	US-08-469-461-3	Sequence 3, Appl1
C 454	15	1.0	3000	6	US-08-680-395-1	Sequence 1, Appl1	C 527	15	1.0	22846	5	US-07-890-609-3	Sequence 3, Appl1
C 455	15	1.0	3000	6	PCT-US95-00601-1	Sequence 1, Appl1	C 528	15	1.0	28994	5	US-08-884-324-14	Sequence 14, Appl1
C 456	15	1.0	3060	5	US-08-098-141-1	Sequence 1, Appl1	C 529	15	1.0	246240	3	US-08-724-394A-20	Sequence 20, Appl1
C 457	15	1.0	3060	5	US-08-968-752B-3	Sequence 3, Appl1	C 530	15	1.0	246240	3	US-08-724-394A-22	Sequence 22, Appl1
C 458	15	1.0	3167	4	US-08-472-534-1	Sequence 1, Appl1	C 531	15	1.0	246240	3	US-08-724-394A-22	Sequence 22, Appl1
C 459	15	1.0	3222	5	US-08-968-752B-1	Sequence 1, Appl1	C 532	15	1.0	18	5	US-08-411-796-403	Sequence 403, App
C 460	15	1.0	3222	5	US-07-906-349A-1	Sequence 1, Appl1	C 533	15	1.0	18	5	US-08-411-796-403	Sequence 403, App
C 461	15	1.0	3372	1	US-08-167-035-1	Sequence 1, Appl1	C 534	15	1.0	18	5	US-08-411-796-403	Sequence 403, App
C 462	15	1.0	3372	1	US-08-167-035-48	Sequence 48, Appl1	C 535	15	1.0	21	6	PCT-US93-11198-403	Sequence 403, App
C 463	15	1.0	3372	1	US-08-208-887A-1	Sequence 1, Appl1	C 536	15	1.0	21	6	US-08-457-273B-40	Sequence 40, Appl1
C 464	15	1.0	3372	3	US-08-539-005-1	Sequence 1, Appl1	C 537	15	1.0	21	3	US-08-450-905B-146	Sequence 146, App

C 538	14	0.9	21	5	US-07-982-759F-146	Sequence 146, App	C 611	14	0.9	480	2	US-07-847-010-10	Sequence 10, App
539	14	0.9	24	4	US-08-913-477-5	Sequence 5, App	C 612	14	0.9	480	2	US-07-847-010-13	Sequence 13, App
540	14	0.9	24	4	US-08-913-477-5	Sequence 24, App	C 613	14	0.9	480	2	US-07-847-010-16	Sequence 16, App
541	14	0.9	27	1	US-08-220-606B-38	Sequence 38, App	C 614	14	0.9	480	2	US-07-847-010-19	Sequence 19, App
542	14	0.9	27	3	US-08-599-602-10	Sequence 10, App	C 615	14	0.9	483	2	US-08-592-126-118	Sequence 118, App
C 543	14	0.9	29	1	US-08-246-982A-14	Sequence 14, App	C 616	14	0.9	516	5	US-08-784-551C-1	Sequence 1, App
C 544	14	0.9	29	1	US-08-453-265-14	Sequence 14, App	C 617	14	0.9	525	1	US-08-434-411-1	Sequence 1, App
545	14	0.9	32	3	US-08-896-410-33	Sequence 33, App	C 618	14	0.9	525	2	US-08-434-402-1	Sequence 1, App
546	14	0.9	33	3	US-08-468-558-18	Sequence 18, App	C 619	14	0.9	525	2	US-08-783-288-1	Sequence 1, App
C 547	14	0.9	35	3	US-08-468-558-17	Sequence 17, App	C 620	14	0.9	525	4	US-08-890-640-1	Sequence 1, App
548	14	0.9	38	3	US-08-292-620A-2282	Sequence 2282, Ap	C 621	14	0.9	525	6	PCT-US94-12873-1	Sequence 1, App
C 549	14	0.9	42	2	US-08-261-660A-28	Sequence 28, App	C 622	14	0.9	525	7	5194592-25	Sequence 1, App
C 550	14	0.9	42	6	PCT-US94-06931-28	Sequence 28, App	C 623	14	0.9	537	6	PCT-US94-18873-16	Sequence 16, App
C 551	14	0.9	45	2	US-08-661-507-1	Sequence 1, App	C 624	14	0.9	541	3	US-08-787-689-13	Sequence 13, App
C 552	14	0.9	45	3	US-08-855-085-2	Sequence 2, App	C 625	14	0.9	546	5	US-08-469-118-177	Sequence 177, App
C 553	14	0.9	45	4	US-09-186-030-2	Sequence 2, App	C 626	14	0.9	546	5	US-08-469-118-178	Sequence 178, App
C 554	14	0.9	45	4	US-08-865-675-2	Sequence 2, App	C 627	14	0.9	546	5	US-08-469-118-179	Sequence 179, App
C 555	14	0.9	45	4	US-08-933-749-3	Sequence 3, App	C 628	14	0.9	546	6	PCT-US95-01185-177	Sequence 178, App
C 556	14	0.9	45	4	US-09-237-510-2	Sequence 2, App	C 629	14	0.9	546	6	PCT-US95-01185-178	Sequence 179, App
C 557	14	0.9	45	5	US-09-120-916-2	Sequence 2, App	C 630	14	0.9	546	6	PCT-US95-01185-179	Sequence 179, App
C 558	14	0.9	45	5	US-08-964-020-6	Sequence 6, App	C 631	14	0.9	555	6	PCT-US94-12873-13	Sequence 13, App
C 559	14	0.9	46	2	US-08-358-344B-8	Sequence 8, App	C 632	14	0.9	555	6	PCT-US94-12873-13	Sequence 13, App
C 560	14	0.9	46	2	US-08-358-344B-10	Sequence 10, App	C 633	14	0.9	555	6	PCT-US94-12873-13	Sequence 13, App
561	14	0.9	49	2	US-08-261-660A-27	Sequence 27, App	C 634	14	0.9	597	5	US-08-742-023-3	Sequence 3, App
C 562	14	0.9	49	3	US-08-652-558-20	Sequence 20, App	C 635	14	0.9	597	5	US-08-968-505-3	Sequence 3, App
C 563	14	0.9	49	3	US-08-652-558-20	Sequence 20, App	C 636	14	0.9	629	2	US-07-847-010-9	Sequence 9, App
C 564	14	0.9	49	3	US-08-652-558-21	Sequence 21, App	C 637	14	0.9	629	2	US-08-592-126-135	Sequence 135, App
565	14	0.9	49	6	PCT-US94-06931-27	Sequence 27, App	C 638	14	0.9	655	2	US-07-847-010-15	Sequence 15, App
566	14	0.9	56	1	US-08-197-770-19	Sequence 19, App	C 639	14	0.9	677	6	PCT-US94-10973A-27	Sequence 27, App
C 567	14	0.9	56	1	US-08-049-264C-15	Sequence 15, App	C 640	14	0.9	692	2	US-08-465-293A-1	Sequence 1, App
C 568	14	0.9	56	1	US-08-476-562-15	Sequence 15, App	C 641	14	0.9	692	2	US-08-465-293A-1	Sequence 1, App
C 569	14	0.9	56	2	US-08-479-723A-15	Sequence 15, App	C 642	14	0.9	728	5	US-09-102-977-1	Sequence 1, App
C 570	14	0.9	56	6	PCT-US94-04310-15	Sequence 15, App	C 643	14	0.9	728	5	US-08-718-804-4	Sequence 4, App
C 571	14	0.9	58	7	5252474-1	Patent No. 5252474	C 644	14	0.9	728	6	PCT-US95-10973A-28	Sequence 28, App
C 572	14	0.9	72	3	US-08-599-602-5	Sequence 5, App	C 645	14	0.9	742	2	US-07-847-010-12	Sequence 12, App
C 573	14	0.9	120	4	US-08-484-575A-23	Sequence 23, App	C 646	14	0.9	750	4	US-08-687-080-74	Sequence 74, App
574	14	0.9	129	3	US-08-599-602-3	Sequence 3, App	C 647	14	0.9	777	3	US-08-618-911-3	Sequence 3, App
575	14	0.9	160	6	PCT-US93-06251-32	Sequence 32, App	C 648	14	0.9	777	3	US-08-618-911-5	Sequence 5, App
C 576	14	0.9	188	4	US-08-928-682-39	Sequence 39, App	C 649	14	0.9	795	3	US-08-898-779-1	Sequence 1, App
C 577	14	0.9	200	6	PCT-US94-09450-10	Sequence 10, App	C 650	14	0.9	795	3	US-08-898-779-3	Sequence 3, App
C 578	14	0.9	225	1	US-08-264-534-14	Sequence 14, App	C 651	14	0.9	816	4	US-08-822-028-26	Sequence 26, App
C 579	14	0.9	225	2	US-08-465-500-24	Sequence 24, App	C 652	14	0.9	816	4	US-08-822-028-25	Sequence 25, App
C 580	14	0.9	225	3	US-08-346-126-24	Sequence 24, App	C 653	14	0.9	830	2	US-08-822-028-25	Sequence 25, App
C 581	14	0.9	225	3	US-08-346-126-24	Sequence 24, App	C 654	14	0.9	853	1	US-07-876-284-1	Sequence 1, App
C 582	14	0.9	255	2	US-08-680-726A-83	Sequence 83, App	C 655	14	0.9	860	2	US-07-847-010-18	Sequence 18, App
C 583	14	0.9	295	1	US-08-264-534-15	Sequence 15, App	C 656	14	0.9	864	4	US-08-924-759-15	Sequence 15, App
C 584	14	0.9	295	2	US-08-083-590A-7	Sequence 7, App	C 657	14	0.9	864	4	US-08-822-028-5	Sequence 5, App
C 585	14	0.9	295	2	US-08-465-500-15	Sequence 15, App	C 658	14	0.9	864	4	US-08-595-668C-11	Sequence 11, App
C 586	14	0.9	295	3	US-08-346-126-15	Sequence 15, App	C 659	14	0.9	865	4	US-08-822-028-3	Sequence 3, App
C 587	14	0.9	318	1	US-08-346-126-15	Sequence 15, App	C 660	14	0.9	871	6	PCT-US91-06234-7	Sequence 7, App
588	14	0.9	318	1	US-08-171-385-3	Sequence 3, App	C 661	14	0.9	906	5	US-08-469-118-53	Sequence 53, App
589	14	0.9	318	5	US-08-361-441B-3	Sequence 3, App	C 662	14	0.9	906	5	PCT-US95-01185-53	Sequence 53, App
590	14	0.9	400	5	US-08-714-918-23	Sequence 23, App	C 663	14	0.9	921	5	US-08-469-118-56	Sequence 56, App
591	14	0.9	423	1	US-08-470-179-77	Sequence 77, App	C 664	14	0.9	921	5	US-08-469-118-60	Sequence 60, App
C 592	14	0.9	425	1	US-08-470-179-134	Sequence 134, App	C 665	14	0.9	921	5	US-08-469-118-60	Sequence 60, App
C 593	14	0.9	444	2	US-08-688-609-2	Sequence 2, App	C 666	14	0.9	921	5	US-08-469-118-72	Sequence 72, App
594	14	0.9	461	3	US-08-977-554-3	Sequence 3, App	C 667	14	0.9	921	5	US-08-469-118-75	Sequence 75, App
595	14	0.9	470	1	US-07-998-003A-66	Sequence 66, App	C 668	14	0.9	921	5	US-08-469-118-80	Sequence 80, App
596	14	0.9	470	1	US-08-453-274B-66	Sequence 66, App	C 669	14	0.9	921	5	US-08-469-118-84	Sequence 84, App
597	14	0.9	470	2	US-08-453-695A-66	Sequence 66, App	C 670	14	0.9	921	5	US-08-469-118-85	Sequence 85, App
598	14	0.9	470	2	US-08-268-161A-66	Sequence 66, App	C 671	14	0.9	921	5	US-08-469-118-85	Sequence 85, App
599	14	0.9	470	2	US-08-453-702A-66	Sequence 66, App	C 672	14	0.9	921	5	US-08-469-118-85	Sequence 85, App
600	14	0.9	470	6	PCT-US93-12588-66	Sequence 66, App	C 673	14	0.9	921	5	US-08-469-118-85	Sequence 85, App
601	14	0.9	470	6	PCT-US93-12588-66	Sequence 66, App	C 674	14	0.9	921	5	PCT-US93-01185-56	Sequence 56, App
602	14	0.9	470	6	PCT-US93-08071-66	Sequence 66, App	C 675	14	0.9	921	6	PCT-US95-01185-59	Sequence 59, App
603	14	0.9	471	3	US-08-599-602-1	Sequence 1, App	C 676	14	0.9	921	6	PCT-US95-01185-60	Sequence 60, App
604	14	0.9	473	2	US-08-480-784-23	Sequence 23, App	C 677	14	0.9	921	6	PCT-US95-01185-72	Sequence 72, App
605	14	0.9	473	2	US-08-483-553-23	Sequence 23, App	C 678	14	0.9	921	6	PCT-US93-01185-75	Sequence 75, App
606	14	0.9	473	2	US-08-487-002-23	Sequence 23, App	C 679	14	0.9	921	6	PCT-US93-01185-78	Sequence 78, App
607	14	0.9	473	2	US-08-483-554B-23	Sequence 23, App	C 680	14	0.9	921	6	PCT-US95-01185-84	Sequence 84, App
608	14	0.9	473	2	US-08-488-011B-23	Sequence 23, App	C 681	14	0.9	921	6	PCT-US95-01185-85	Sequence 85, App
609	14	0.9	473	6	PCT-US95-10202-23	Sequence 23, App	C 682	14	0.9	921	6	PCT-US95-01185-85	Sequence 85, App
610	14	0.9	473	6	PCT-US95-10220-23	Sequence 23, App	C 683	14	0.9	924	1	US-07-593-657-4	Sequence 4, App

C 684	14	0.9	950	5	US-08-581-148C-10	Sequence 10, Appl	757	14	0.9	1299	2	US-08-463-262A-1	Sequence 1, Appl1
C 685	14	0.9	966	5	US-08-469-318-68	Sequence 68, Appl	758	14	0.9	1299	2	US-08-463-989-1	Sequence 1, Appl1
C 686	14	0.9	966	5	US-08-469-318-70	Sequence 70, Appl	C 759	14	0.9	1301	6	PCT-US96-1051-19	Sequence 19, Appl
C 687	14	0.9	966	5	US-08-469-318-71	Sequence 71, Appl	C 760	14	0.9	1309	4	US-08-933-750C-63	Sequence 63, Appl
C 688	14	0.9	966	5	US-08-469-318-73	Sequence 73, Appl	C 761	14	0.9	1329	1	US-08-278-630A-8	Sequence 8, Appl
C 689	14	0.9	966	5	US-08-469-318-77	Sequence 77, Appl	C 762	14	0.9	1330	3	US-08-807-200-1	Sequence 1, Appl1
C 690	14	0.9	966	5	US-08-469-318-79	Sequence 79, Appl	C 763	14	0.9	1345	4	US-08-702-153-3	Sequence 3, Appl1
C 691	14	0.9	966	5	US-08-469-318-81	Sequence 81, Appl	C 764	14	0.9	1350	1	US-07-792-466-7	Sequence 7, Appl1
C 692	14	0.9	966	5	US-08-469-318-89	Sequence 89, Appl	C 765	14	0.9	1350	2	US-08-474-633A-1	Sequence 1, Appl1
C 693	14	0.9	966	6	PCT-US95-01185-68	Sequence 68, Appl	C 766	14	0.9	1350	6	US-08-737-552B-4	Sequence 4, Appl1
C 694	14	0.9	966	6	PCT-US95-01185-70	Sequence 70, Appl	C 767	14	0.9	1350	6	PCT-US93-02480-1	Sequence 1, Appl1
C 695	14	0.9	966	6	PCT-US95-01185-71	Sequence 71, Appl	C 768	14	0.9	1355	3	US-08-618-408B-3	Sequence 3, Appl1
C 696	14	0.9	966	6	PCT-US95-01185-73	Sequence 73, Appl	C 769	14	0.9	1356	3	US-08-484-126-4	Sequence 4, Appl1
C 697	14	0.9	966	6	PCT-US95-01185-77	Sequence 77, Appl	C 770	14	0.9	1373	6	PCT-US92-02977-6	Sequence 6, Appl1
C 698	14	0.9	966	6	PCT-US95-01185-79	Sequence 79, Appl	C 771	14	0.9	1373	6	PCT-US95-03032-2	Sequence 3, Appl1
C 699	14	0.9	966	6	PCT-US95-01185-81	Sequence 81, Appl	C 772	14	0.9	1383	1	US-08-289-709-2	Sequence 2, Appl1
C 700	14	0.9	966	6	PCT-US95-01185-89	Sequence 89, Appl	C 773	14	0.9	1383	1	US-08-602-656-2	Sequence 2, Appl1
C 701	14	0.9	982	4	US-08-822-028-1	Sequence 1, Appl1	C 774	14	0.9	1389	2	US-08-702-344-20	Sequence 20, Appl
C 702	14	0.9	984	5	US-08-469-318-83	Sequence 83, Appl	C 775	14	0.9	1395	1	US-07-991-867B-25	Sequence 25, Appl
C 703	14	0.9	984	5	PCT-US95-01185-83	Sequence 83, Appl	C 776	14	0.9	1395	1	US-07-991-867B-25	Sequence 25, Appl
C 704	14	0.9	1005	5	US-09-103-359-4	Sequence 4, Appl1	C 777	14	0.9	1395	2	US-08-107-755A-25	Sequence 25, Appl
C 705	14	0.9	1017	5	US-08-469-318-67	Sequence 67, Appl	C 778	14	0.9	1395	2	US-08-107-755A-25	Sequence 25, Appl
C 706	14	0.9	1017	5	PCT-US95-01185-67	Sequence 67, Appl	C 779	14	0.9	1395	4	US-08-544-332-25	Sequence 25, Appl
C 707	14	0.9	1018	7	5187262-7	Sequence 67, Appl	C 780	14	0.9	1395	4	US-08-544-332-25	Sequence 25, Appl
C 708	14	0.9	1047	5	US-08-469-318-65	Sequence 65, Appl	C 781	14	0.9	1400	3	US-08-544-332-25	Sequence 25, Appl
C 709	14	0.9	1047	5	US-08-469-318-74	Sequence 74, Appl	C 782	14	0.9	1400	3	US-08-305-764C-59	Sequence 59, Appl
C 710	14	0.9	1047	5	US-08-469-318-76	Sequence 76, Appl	C 783	14	0.9	1425	1	US-07-715-184-3	Sequence 3, Appl1
C 711	14	0.9	1047	6	PCT-US95-01185-65	Sequence 65, Appl	C 784	14	0.9	1425	1	US-07-876-280-6	Sequence 6, Appl1
C 712	14	0.9	1047	6	PCT-US95-01185-74	Sequence 74, Appl	C 785	14	0.9	1425	1	US-07-876-280-6	Sequence 6, Appl1
C 713	14	0.9	1047	6	PCT-US95-01185-76	Sequence 76, Appl	C 786	14	0.9	1425	1	US-07-935-310A-1	Sequence 1, Appl1
C 714	14	0.9	1057	2	US-08-470-202-45	Sequence 45, Appl	C 787	14	0.9	1425	1	US-07-828-788A-1	Sequence 1, Appl1
C 715	14	0.9	1057	2	US-08-471-770-45	Sequence 45, Appl	C 788	14	0.9	1425	1	US-08-049-783-3	Sequence 3, Appl1
C 716	14	0.9	1057	2	US-08-468-059-45	Sequence 45, Appl	C 789	14	0.9	1425	1	US-08-147-189-1	Sequence 1, Appl1
C 717	14	0.9	1062	1	US-08-426-169-6	Sequence 6, Appl1	C 790	14	0.9	1425	2	US-08-316-301A-7	Sequence 2, Appl1
C 718	14	0.9	1062	6	PCT-US95-09470-6	Sequence 6, Appl1	C 791	14	0.9	1425	3	US-08-904-278-1	Sequence 3, Appl1
C 719	14	0.9	1080	7	5474933-1	Sequence 6, Appl1	C 792	14	0.9	1425	6	PCT-US92-03624-7	Sequence 7, Appl1
C 720	14	0.9	1083	6	PCT-US94-09450-7	Sequence 7, Appl1	C 793	14	0.9	1425	6	PCT-US92-11337-1	Sequence 1, Appl1
C 721	14	0.9	1121	2	US-08-523-855A-2	Sequence 2, Appl1	C 794	14	0.9	1425	6	PCT-US93-07409-1	Sequence 1, Appl1
C 722	14	0.9	1121	7	5256558-7	Sequence 7, Appl1	C 795	14	0.9	1434	3	US-08-476-254-1	Sequence 3, Appl1
C 723	14	0.9	1124	3	US-08-578-551-1	Sequence 1, Appl1	C 796	14	0.9	1440	5	US-08-807-200-11	Sequence 11, Appl
C 724	14	0.9	1124	4	US-09-190-982-1	Sequence 1, Appl1	C 797	14	0.9	1440	5	US-08-852-782A-2	Sequence 2, Appl1
C 725	14	0.9	1126	6	PCT-US95-17083-16	Sequence 16, Appl	C 798	14	0.9	1443	6	PCT-US96-10521-13	Sequence 3, Appl
C 726	14	0.9	1182	3	US-08-977-554-1	Sequence 1, Appl1	C 799	14	0.9	1463	6	PCT-US93-05640-30	Sequence 10, Appl
C 727	14	0.9	1188	3	US-08-921-426-6	Sequence 6, Appl1	C 800	14	0.9	1499	1	US-08-415-818-11	Sequence 11, Appl
C 728	14	0.9	1188	5	US-08-816-915-6	Sequence 6, Appl1	C 801	14	0.9	1499	4	US-08-894-236-11	Sequence 11, Appl
C 729	14	0.9	1188	6	PCT-US95-07743-6	Sequence 6, Appl1	C 802	14	0.9	1499	6	PCT-US96-01444-11	Sequence 11, Appl
C 730	14	0.9	1189	4	US-08-450-042A-2	Sequence 2, Appl1	C 803	14	0.9	1506	1	US-08-149-105-2	Sequence 1, Appl
C 731	14	0.9	1195	7	5240848-6	Sequence 2, Appl1	C 804	14	0.9	1506	1	US-08-317-847-2	Sequence 2, Appl1
C 732	14	0.9	1200	3	US-08-642-045B-3	Sequence 3, Appl1	C 805	14	0.9	1512	2	US-08-594-808B-6	Sequence 6, Appl1
C 733	14	0.9	1201	1	US-08-181-271A-26	Sequence 26, Appl	C 806	14	0.9	1524	2	US-08-409-122-1	Sequence 1, Appl1
C 734	14	0.9	1201	1	US-08-449-315-26	Sequence 26, Appl	C 807	14	0.9	1524	2	US-08-408-666-1	Sequence 1, Appl1
C 735	14	0.9	1201	1	US-08-444-803-26	Sequence 26, Appl	C 808	14	0.9	1541	3	US-08-668-289-9	Sequence 9, Appl1
C 736	14	0.9	1201	1	US-08-449-043-26	Sequence 26, Appl	C 809	14	0.9	1541	3	US-09-167-112-9	Sequence 9, Appl1
C 737	14	0.9	1201	2	US-08-456-265A-26	Sequence 26, Appl	C 810	14	0.9	1550	5	US-08-912-560-1	Sequence 1, Appl1
C 738	14	0.9	1201	2	US-08-455-416-26	Sequence 26, Appl	C 811	14	0.9	1552	5	US-08-948-705-1	Sequence 1, Appl1
C 739	14	0.9	1201	2	US-08-455-244-26	Sequence 26, Appl	C 812	14	0.9	1581	4	US-08-853-652A-15	Sequence 15, Appl
C 740	14	0.9	1201	2	US-08-454-876-26	Sequence 26, Appl	C 813	14	0.9	1592	1	US-08-044-618-4	Sequence 4, Appl1
C 741	14	0.9	1201	3	US-08-457-364-26	Sequence 26, Appl	C 814	14	0.9	1593	1	US-08-307-499-50	Sequence 50, Appl
C 742	14	0.9	1201	3	US-08-456-262-26	Sequence 26, Appl	C 815	14	0.9	1603	1	US-07-885-970A-19	Sequence 19, Appl
C 743	14	0.9	1201	3	US-08-456-240-26	Sequence 26, Appl	C 816	14	0.9	1603	1	US-08-298-687A-19	Sequence 19, Appl
C 744	14	0.9	1201	3	US-08-455-736-26	Sequence 26, Appl	C 817	14	0.9	1603	1	US-08-298-829-19	Sequence 19, Appl
C 745	14	0.9	1201	4	US-08-971-217-26	Sequence 26, Appl	C 818	14	0.9	1644	5	US-08-852-782A-1	Sequence 1, Appl1
C 746	14	0.9	1216	4	US-08-078-311-20	Sequence 20, Appl	C 819	14	0.9	1674	2	US-08-523-855A-1	Sequence 1, Appl1
C 747	14	0.9	1216	4	US-08-460-402-20	Sequence 20, Appl	C 820	14	0.9	1686	1	US-08-324-483-1	Sequence 1, Appl1
C 748	14	0.9	1230	2	US-08-700-442A-8	Sequence 8, Appl1	C 821	14	0.9	1732	2	US-08-480-346-2	Sequence 2, Appl1
C 749	14	0.9	1242	2	US-07-877-516-1	Sequence 1, Appl1	C 822	14	0.9	1732	3	US-08-243-541-2	Sequence 2, Appl1
C 750	14	0.9	1242	6	PCT-US93-03077-4	Sequence 4, Appl1	C 823	14	0.9	1732	3	US-08-480-344-2	Sequence 2, Appl1
C 751	14	0.9	1248	2	US-08-602-010A-9	Sequence 9, Appl1	C 824	14	0.9	1743	1	US-08-171-382-3	Sequence 3, Appl1
C 752	14	0.9	1248	2	US-08-680-726A-9	Sequence 9, Appl1	C 825	14	0.9	1743	1	US-08-309-420-3	Sequence 3, Appl1
C 753	14	0.9	1263	6	PCT-US96-10602-9	Sequence 9, Appl1	C 826	14	0.9	1743	1	US-08-309-419-3	Sequence 3, Appl1
C 754	14	0.9	1278	4	US-08-765-268A-2	Sequence 2, Appl1	C 827	14	0.9	1743	3	US-08-841-178-20	Sequence 20, Appl
C 755	14	0.9	1286	3	US-08-807-200-3	Sequence 3, Appl1	C 828	14	0.9	1743	6	PCT-US95-11856-3	Sequence 3, Appl1
C 756	14	0.9	1294	2	US-08-464-523B-3	Sequence 3, Appl1	C 829	14	0.9	1743	6	PCT-US95-11878-3	Sequence 3, Appl1

C 830	14	0.9	1752	1	US-07-629-1041-1	Sequence 1, Appl1	C 903	14	0.9	2278	2	US-08-526-813-1	Sequence 1, Appl1
C 831	14	0.9	1767	1	US-08-399-646-1	Sequence 1, Appl1	C 904	14	0.9	2278	6	PCT-US95-08554-1	Sequence 1, Appl1
C 832	14	0.9	1767	2	US-08-607-321-1	Sequence 1, Appl1	C 905	14	0.9	2314	5	US-08-105-444-1	Sequence 1, Appl1
C 833	14	0.9	1767	3	US-08-961-240-1	Sequence 1, Appl1	C 906	14	0.9	2322	1	US-08-618-164-1	Sequence 1, Appl1
C 834	14	0.9	1767	3	US-08-605-501-1	Sequence 1, Appl1	C 907	14	0.9	2322	2	US-08-314-309A-20	Sequence 20, Appl1
C 835	14	0.9	1769	2	US-08-233-155A-19	Sequence 19, Appl1	C 908	14	0.9	2336	4	US-08-389-564B-2	Sequence 2, Appl1
C 836	14	0.9	1769	4	US-08-765-268A-1	Sequence 1, Appl1	C 909	14	0.9	2336	5	US-08-466-047B-2	Sequence 2, Appl1
C 837	14	0.9	1776	2	US-08-464-523B-4	Sequence 4, Appl1	C 910	14	0.9	2363	4	US-08-671-978A-11	Sequence 11, Appl1
C 838	14	0.9	1778	6	PCT-US91-02958-9	Sequence 9, Appl1	C 911	14	0.9	2382	1	US-08-256-938-1	Sequence 1, Appl1
C 839	14	0.9	1780	5	PCT-US94-12913A-17	Sequence 17, Appl1	C 912	14	0.9	2392	1	US-08-171-382-5	Sequence 5, Appl1
C 840	14	0.9	1789	5	US-08-948-705-6	Sequence 6, Appl1	C 913	14	0.9	2392	1	US-08-309-420-5	Sequence 5, Appl1
C 841	14	0.9	1823	2	US-08-680-726A-77	Sequence 77, Appl1	C 914	14	0.9	2392	1	US-08-309-419-5	Sequence 5, Appl1
C 842	14	0.9	1823	3	US-08-680-726A-78	Sequence 78, Appl1	C 915	14	0.9	2392	6	PCT-US95-11856-5	Sequence 5, Appl1
C 843	14	0.9	1840	3	US-08-484-993B-15	Sequence 15, Appl1	C 916	14	0.9	2392	6	PCT-US95-11878-5	Sequence 5, Appl1
C 844	14	0.9	1840	4	US-08-484-158B-15	Sequence 15, Appl1	C 917	14	0.9	2401	2	US-08-480-346-1	Sequence 1, Appl1
C 845	14	0.9	1840	4	US-08-484-596A-15	Sequence 15, Appl1	C 918	14	0.9	2401	2	US-08-244-541-1	Sequence 1, Appl1
C 846	14	0.9	1840	4	US-08-480-150A-15	Sequence 15, Appl1	C 919	14	0.9	2401	3	US-08-480-344-1	Sequence 1, Appl1
C 847	14	0.9	1840	5	US-08-458-731-15	Sequence 15, Appl1	C 920	14	0.9	2409	1	US-07-911-531-18	Sequence 18, Appl1
C 848	14	0.9	1840	5	US-08-149-223A-15	Sequence 15, Appl1	C 921	14	0.9	2409	1	US-07-693-636A-18	Sequence 18, Appl1
C 849	14	0.9	1867	3	US-08-516-801-1	Sequence 1, Appl1	C 922	14	0.9	2417	2	US-08-011-398B-1	Sequence 1, Appl1
C 850	14	0.9	1867	6	PCT-US95-06683-1	Sequence 49, Appl1	C 923	14	0.9	2417	2	US-08-464-051-1	Sequence 1, Appl1
C 851	14	0.9	1880	1	US-08-247-475-49	Sequence 49, Appl1	C 924	14	0.9	2417	3	US-08-462-498-1	Sequence 1, Appl1
C 852	14	0.9	1880	1	US-08-479-650-48	Sequence 49, Appl1	C 925	14	0.9	2417	3	US-08-554-385-2	Sequence 2, Appl1
C 853	14	0.9	1880	2	US-08-191-866D-80	Sequence 80, Appl1	C 926	14	0.9	2427	2	US-08-490-099-1	Sequence 1, Appl1
C 854	14	0.9	1880	2	US-08-674-169-49	Sequence 49, Appl1	C 927	14	0.9	2430	3	US-08-062-368-1	Sequence 35, Appl1
C 855	14	0.9	1880	3	US-08-185-949B-80	Sequence 80, Appl1	C 928	14	0.9	2430	3	US-08-820-170A-35	Sequence 16, Appl1
C 856	14	0.9	1883	2	US-08-665-220-3	Sequence 3, Appl1	C 929	14	0.9	2436	2	US-08-306-601B-16	Sequence 16, Appl1
C 857	14	0.9	1911	2	US-08-258-420-7	Sequence 8, Appl1	C 930	14	0.9	2448	4	US-08-687-080-111	Sequence 11, App
C 858	14	0.9	1914	2	US-08-258-420-7	Sequence 7, Appl1	C 931	14	0.9	2451	3	US-08-820-170A-29	Sequence 29, Appl1
C 859	14	0.9	1930	1	US-07-982-112-1	Sequence 1, Appl1	C 932	14	0.9	2455	1	US-08-256-938-3	Sequence 3, Appl1
C 860	14	0.9	1939	7	US-07-982-112-1	Patent No. 5198542-3	C 933	14	0.9	2455	1	US-08-797-689-15	Sequence 15, Appl1
C 861	14	0.9	1942	5	US-08-627-907A-3	Sequence 3, Appl1	C 934	14	0.9	2471	1	US-08-444-231-18	Sequence 18, Appl1
C 862	14	0.9	2009	2	US-08-201-118-6	Sequence 6, Appl1	C 935	14	0.9	2471	1	US-08-152-443A-18	Sequence 18, Appl1
C 863	14	0.9	2009	4	US-08-238-821B-6	Sequence 6, Appl1	C 936	14	0.9	2472	2	US-08-425-289A-2	Sequence 2, Appl1
C 864	14	0.9	2009	6	PCT-US93-05744-6	Sequence 6, Appl1	C 937	14	0.9	2474	4	US-08-666-352A-2	Sequence 2, Appl1
C 865	14	0.9	2014	3	US-08-525-742-7	Sequence 7, Appl1	C 938	14	0.9	2487	3	US-08-820-170A-32	Sequence 32, Appl1
C 866	14	0.9	2020	3	US-08-417-210A-136	Sequence 136, App	C 939	14	0.9	2505	3	US-08-890-094-1	Sequence 1, Appl1
C 867	14	0.9	2030	4	US-08-923-536A-1	Sequence 1, Appl1	C 940	14	0.9	2517	1	US-07-906-930E-1	Sequence 1, Appl1
C 868	14	0.9	2049	3	US-08-417-210A-117	Sequence 117, App	C 941	14	0.9	2529	3	US-08-867-129-1	Sequence 1, Appl1
C 869	14	0.9	2070	3	US-08-890-094-15	Sequence 17, Appl1	C 942	14	0.9	2534	3	US-08-219-237B-1	Sequence 1, Appl1
C 870	14	0.9	2071	2	US-08-393-985-15	Sequence 15, Appl1	C 943	14	0.9	2550	7	US-08-219-237B-1	Sequence 1, Appl1
C 871	14	0.9	2074	1	US-08-414-685-1	Sequence 1, Appl1	C 944	14	0.9	2559	3	US-08-248-839C-10	Sequence 10, Appl1
C 872	14	0.9	2095	1	US-08-261-086-1	Sequence 1, Appl1	C 945	14	0.9	2587	3	US-08-326-286-3	Sequence 3, Appl1
C 873	14	0.9	2095	1	US-08-305-764C-55	Sequence 55, Appl1	C 946	14	0.9	2619	6	PCT-US96-10521-17	Sequence 17, Appl1
C 874	14	0.9	2115	1	US-08-329-681A-3	Sequence 3, Appl1	C 947	14	0.9	2629	4	US-09-012-030-1	Sequence 1, Appl1
C 875	14	0.9	2126	2	US-08-463-262A-2	Sequence 2, Appl1	C 948	14	0.9	2629	4	US-08-590-454-1	Sequence 1, Appl1
C 876	14	0.9	2126	2	US-08-463-985-2	Sequence 2, Appl1	C 949	14	0.9	2631	2	US-08-208-036-13	Sequence 13, Appl1
C 877	14	0.9	2147	1	US-08-256-136-1	Sequence 1, Appl1	C 950	14	0.9	2631	2	US-08-428-823-13	Sequence 13, Appl1
C 878	14	0.9	2147	4	US-08-950-737-1	Sequence 7, Appl1	C 951	14	0.9	2649	4	US-08-718-964-1	Sequence 1, Appl1
C 879	14	0.9	2147	4	US-08-973-461A-7	Sequence 7, Appl1	C 952	14	0.9	2649	4	US-09-059-964A-1	Sequence 1, Appl1
C 880	14	0.9	2147	5	US-08-648-010-7	Sequence 7, Appl1	C 953	14	0.9	2649	4	US-08-842-341-1	Sequence 1, Appl1
C 881	14	0.9	2150	1	US-08-198-446B-10	Sequence 10, Appl1	C 954	14	0.9	2661	2	US-08-351-413-1	Sequence 1, Appl1
C 882	14	0.9	2150	3	US-08-870-693-10	Sequence 10, Appl1	C 955	14	0.9	2682	1	US-09-025-583-1	Sequence 1, Appl1
C 883	14	0.9	2152	1	US-07-779-890-1	Sequence 1, Appl1	C 956	14	0.9	2682	4	US-08-044-618-3	Sequence 3, Appl1
C 884	14	0.9	2152	2	US-07-779-890-1	Sequence 1, Appl1	C 957	14	0.9	2682	4	US-08-867-941-7	Sequence 7, Appl1
C 885	14	0.9	2152	6	PCT-US93-05640-1	Sequence 1, Appl1	C 958	14	0.9	2689	4	US-08-876-546A-15	Sequence 15, Appl1
C 886	14	0.9	2152	6	PCT-US93-05640-1	Sequence 11, Appl1	C 959	14	0.9	2692	1	US-07-932-454A-2	Sequence 2, Appl1
C 887	14	0.9	2161	2	US-08-607-321-11	Sequence 11, Appl1	C 960	14	0.9	2694	1	US-08-867-941-2	Sequence 2, Appl1
C 888	14	0.9	2161	3	US-08-605-501-11	Sequence 11, Appl1	C 961	14	0.9	2733	2	US-08-676-967-3	Sequence 3, Appl1
C 889	14	0.9	2161	3	US-08-605-501-11	Sequence 11, Appl1	C 962	14	0.9	2733	2	US-08-676-967-3	Sequence 3, Appl1
C 890	14	0.9	2180	5	US-09-003-217-1	Sequence 83, Appl1	C 963	14	0.9	2737	1	US-08-530-010-2	Sequence 2, Appl1
C 891	14	0.9	2185	4	US-08-714-918-83	Sequence 83, Appl1	C 964	14	0.9	2787	1	US-08-530-010-4	Sequence 4, Appl1
C 892	14	0.9	2188	1	US-07-882-925A-4	Sequence 4, Appl1	C 965	14	0.9	2787	1	US-08-530-010-6	Sequence 6, Appl1
C 893	14	0.9	2188	1	US-08-184-012C-4	Sequence 21, Appl1	C 966	14	0.9	2787	1	US-08-530-010-8	Sequence 8, Appl1
C 894	14	0.9	2200	2	US-08-272-235-21	Sequence 21, Appl1	C 967	14	0.9	2787	1	US-08-530-010-10	Sequence 10, Appl1
C 895	14	0.9	2200	6	PCT-US95-08565-21	Sequence 17, Appl1	C 968	14	0.9	2787	2	US-08-484-101B-2	Sequence 2, Appl1
C 896	14	0.9	2219	5	US-08-510-646B-17	Sequence 12, Appl1	C 969	14	0.9	2787	2	US-08-484-101B-6	Sequence 6, Appl1
C 897	14	0.9	2234	5	US-08-714-918-75	Sequence 75, Appl1	C 970	14	0.9	2787	2	US-08-484-101B-8	Sequence 8, Appl1
C 898	14	0.9	2258	4	US-08-201-118-12	Sequence 12, Appl1	C 971	14	0.9	2787	2	US-08-484-101B-8	Sequence 8, Appl1
C 899	14	0.9	2258	4	US-08-238-821B-12	Sequence 12, Appl1	C 972	14	0.9	2787	2	US-08-484-101B-8	Sequence 8, Appl1
C 900	14	0.9	2258	6	PCT-US95-05744-12	Sequence 12, Appl1	C 973	14	0.9	2815	2	US-08-484-101B-10	Sequence 10, Appl1
C 901	14	0.9	2268	3	US-08-873-093-2	Sequence 2, Appl1	C 974	14	0.9	2815	2	US-07-671-817A-2	Sequence 2, Appl1
C 902	14	0.9	2278	1	US-08-258-188-1	Sequence 1, Appl1	C 975	14	0.9	2817	6	PCT-US93-05444-1	Sequence 1, Appl1

Sequence 8, April
Sequence 7, April
Sequence 1, April
Sequence 1, April
Sequence 1, April
Sequence 1, April
Sequence 1, April
Sequence 14, April
Sequence 1, April
Sequence 1, April
Sequence 1, April
Sequence 3, April
Sequence 18, April
Sequence 18, April
Sequence 36, April
Sequence 1, April
Sequence 1, April
Sequence 1, April
Sequence 21, April
Sequence 21, April
Sequence 21, April
Sequence 169, April
Sequence 169, April
Sequence 3, April
Sequence 3, April

RESULT 1
US-08-913-477-1

GENERAL INFORMATION:
APPLICANT: Tilball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.F.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHART P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
City: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913.477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327

```

REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1014 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..987
US-08-913-477-1

```

Query Match 59.2%; Score 927; DB 4; Length 1014;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps

Dy 589 ATTGAGCGCTACGACAACAAACCACACATTTTATTGAGCATCTAGAAAAAGTTAGGGTG 648
 13 ATTGAGCGCTACGACAACAAACCACACATTTTATTGAGCATCTAGAAAAAGTTAGGGTG 72

Dy 649 GAACACTTACTGTGCATCGTTCCTCAGTTTTGAAGAATTGGTTCAGTTAGTCAAGAT 708
 73 GAACACTTACTGTGCATCGTTCCTCAGTTTTGAAGAATTGGTTCAGTTAGTCAAGAT 132

Dy 709 AAAAATATAGATTTTCCATTAAATATGATCCGCAAAAAGATTCCGAGGTTTTGCCAAT 768
Db 133 AAAAATATAGATTTTCCATTAAATATGATCCGCAAAAAGATTCCGAGGTTTTGCCAAT 192

Dy 769 AGAGTAATTACTGATGATATCGAATTGCTCAGAAAAATCCTAGCTTAATTTCTACCCGAG 828
Db 193 AGAGTAATTACTGATGATGATCGAATTGCTCAGAAAAATCCTAGCTTAATTTCTACCCGAG 252

Dy 829 GATACCACTTTTAAAGCGGCTCATTTATGACAACCACTCGAAAATGGCATCAAGGAGTA 888
Db 253 GATGCCATCTTTAAAGCGGCTCATTTATGACAACCACTCGAAAATGGCATCAAGGAGTA 312

Dy 889 AAAGAGTTCCTTGATCATGCGCGAATACCAATGGGAATTCGGGCGTTTCATGGCAGTA 948
Db 313 AAAGAGTTCCTTGATCATGCGCGAATACCAATGGGAATTCGGGCGTTTCATGGCAGTA 372

Dy 949 ATGCATTTCTCTTTAACCGCGCATCGATCGATGATATTTTGAAGTGAATTTGAT 1008
Db 373 ATGCATTTCTCTTTAACCGCGCATCGATCGATGATGATATTTTGAAGTGAATTTGAT 432

Dy 1009 TCAATGATCATCATGTGTATGCCCGTAGCAATTCGGTGAAGAATTTAGCTAGCTTAC 1066
Db 433 TCAATGATCATCATGTGTATGCCCGTAGCAATTCGGTGAAGAATTTAGCTAGCTTAC 492

Dy 1069 GCCGATTTAAAGATTTATAGTTATTCAGAGCGGAATTAATAAGCATGTCGTAAGT 1128
Db 493 GCCGATTTAAAGATTTATAGTTATTCAGAGCGGAATTAATAAGCATGTCGTAAGT 552

Dy 1129 GGCAACATAATATCCATGATTAATCCATTAATCTCATGGATTAATAAATTTATATGGTTAT 1188
Db 553 GGCAACATAATATCCATGATTAATCCATTAATCTCATGGATTAATAAATTTATATGGTTAT 612

Dy 1189 ACAAGTGAAGATTTTAAAGCCAGCGCAGAGTACAAAATTTCTCGAGAAAATGCCCTAA 1248
Db 613 ACAAGTGAAGATTTTAAAGCCAGCGCAGAGTACAAAATTTCTCGAGAAAATGCCCTAA 672

Dy 1249 ACCACCAATTCAGGTGATATGGAGAGCAAAAAAATGCTCGATTAAGAAGACTTTTCTTGA 1300
Db 673 ACCACCAATTCAGGTGATATGGAGAGCAAAAAAATGCTCGATTAAGAAGACTTTTCTTGA 732;

Db 1501 GACACGTCTGTAAATGA 1518

RESULT 4

US-08-913-477-3

Sequence 3, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 No. 5985285th Glade Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913.477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1014 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..987
US-08-913-477-3

Query Match 58.9% Score 923; DB 4; Length 1014;

Best Local Similarity 99.9% Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 973; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
593 GAGCCTAGCAACAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGGTGGAAC 652
|||||

Db 17 GAGCCTAGCAACAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGGTGGAAC 76

Oy 653 AACTACGTGCATGCTCTCTACATTTAGAAATTTGGTTCAGTTACTCAAGATATAA 712

Db 77 AACTACGTGCATGCTCTCTACATTTAGAAATTTGGTTCAGTTACTCAAGATATAA 136

Oy 713 ATATGATATTTCCATTAATATGATCCAGAAAAGATTGGAGGTTTGGCCATAGAG 772

Db 137 ATATGATATTTCCATTAATATGATCCAGAAAAGATTGGAGGTTTGGCCATAGAG 196

Oy 773 TAATACGTGATATGCAATTTGCTCAAGAAAATCCTAGCTATTTTCTACCCGAGATA 832

Db 197 TAATACGTGATATGCAATTTGCTCAAGAAAATCCTAGCTATTTTCTACCCGAGATG 256

Oy 833 CCATCTTAAAGCGCGTATATGCAACCAACTGCCAAAATGGCATCAAGCGATAAAG 892

Db 257 CCATCTTAAAGCGCGTATATGCAACCAACTGCCAAAATGGCATCAAGCGATAAAG 316

Oy 893 AGTTCCTGAATCATCGCCGGAATACAAATGGGAATTTGGGCGCTCATGGCAGTAATGC 952

Db 317 AGTTCCTGAATCATCGCCGGAATACAAATGGGAATTTGGGCGCTCATGGCAGTAATGC 376

Oy 953 ATTCTCTTAAACCGCCGATCGATGATGATATTTTGAAGTATTTGTTGATTCAA 1012

Db 377 ATTCTCTTAAACCGCCGATCGATGATGATATTTTGAAGTATTTGTTGATTCAA 436

Oy 1013 TGAATCATCATGCTATGCTCCGATGACAAAGTTCGGAAGATTTAGCTAGCTACCCCG 1072

Db 437 TGAATCATCATGCTATGCTCCGATGACAAAGTTCGGAAGATTTAGCTAGCTACCCCG 496

Oy 1073 AATTAAGATTTATTCAGTTATTCAGCCGAAATTAATAGATGCTGCTAGTGGCA 1132

Db 497 AATTAAGATTTATTCAGTTATTCAGCCGAAATTAATAGATGCTGCTAGTGGCA 556

Oy 1133 CCATTAATATTCATGATTAATCCATTAATCTCATGATGATTAATTTATATGTTATACG 1192

Db 557 CCATTAATATTCATGATTAATCCATTAATCTCATGATGATTAATTTATATGTTATACG 616

Oy 1193 ATGAAGATTTTAAAGCAGCGAGTACAAAATTTCTCGAGAAATGCCCTCAACCA 1252

Db 617 ATGAAGATTTTAAAGCAGCGAGTACAAAATTTCTCGAGAAATGCCCTCAACCA 676

Oy 1253 CCATCAGGTGATGAGCGGAGAAAATAGTCTCATTAAGAACTTCTGGAAGTG 1312

Db 677 CCATCAGGTGATGAGCGGAGAAAATAGTCTCATTAAGAACTTCTGGAAGTG 736

Oy 1313 AGAATTAAGAAACCGGGCGTTGGTAACTGAAAACCTCATCTATATAAAGATA 1372

Db 737 AGAATTAAGAAACCGGGCGTTGGTAACTGAAAACCTCATCTATATAAAGATA 796

Oy 1373 ATATGATATTTATCTACCTTGGCACCACTGCTGGATAGTCCAGGCCGCTCAAGACT 1432

Db 797 ATATGATATTTATCTACCTTGGCACCACTGCTGGATAGTCCAGGCCGCTCAAGACT 856

Oy 1433 TGGTAGCCAAAACCACTGAGTGTGATTAATACATCACTTTATATACGTATG 1492

Db 857 TGGTAGCCAAAACCACTGAGTGTGATTAATACATCACTTTATATACGTATG 916

Oy 1493 AAGCAGTAAACCGTTTCATTCAGAAATGATTCAGATGCAAGCTGTAGATAGACA 1552

Db 917 AAGCAGTAAACCGTTTCATTCAGAAATGATTCAGATGCAAGCTGTAGATAGACA 976

Oy 1553 CGCTGTGTAATGA 1566

Db 977 CGCTGTGTAATGA 990

RESULT 5

US-08-913-477-20

Sequence 20, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.

Wed Aug 23 11:45:56 2000

us-08-699-716a-1.oliqo.rni

Page 17

```

; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 44::
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-432-871C-44

```

Query Match	1.8%	Score 28	DB 3	Length 39
Best Local Similarity	100.0%	Pred. No. 9	4e-05	
Matches 28	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	19	CATCATCATCATCACACAGCGCGCC	46	
DB	36	CATCATCATCATCACAGCGCGCC	9	

Search completed: August 22, 2000, 16:29:01
Job time: 3138 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 15:10:37 ; Search time 614.39 Seconds
(without alignments)
11240.077 Million cell updates/sec

Title: US-08-699-716A-1

Perfect score: 1566
Sequence: 1 ATGGCCATCATCATCATCA.....ATGACACGCTGTGTAATGA 1566

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5247842 seqs, 2204914090 residues

Word size : 0

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_est56:*
94: gb_est57:*
95: gb_est58:*
96: gb_est59:*
97: em_est61:*
98: em_est62:*
99: em_est63:*
100: em_est64:*
101: gb_est65:*
102: gb_est66:*
103: gb_est67:*
104: gb_est68:*
105: gb_est69:*
106: em_est70:*
107: em_est71:*
108: em_est72:*
109: em_est73:*
110: em_est74:*
111: em_est75:*
112: em_est76:*
113: gb_est77:*
114: gb_est78:*
115: em_est79:*
116: gb_est80:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	33	2.1	480	27	AA542448	AA542448 fa07a10.s
2	33	2.1	595	64	AM154180	AM154180 f123e10.x
3	32	2.0	392	94	AO101925	AO101925 HS_3017.B
4	32	2.0	444	64	AM154525	AM154525 f128b09.x
5	32	2.0	744	117	AO897728	AO897728 HS_3135.A
6	31	2.0	245	106	AO359857	AO359857 RPT-11-3
7	31	2.0	424	79	AM624044	AM624044 EST321989
8	31	2.0	463	95	AO196790	AO196790 CIR-HSP-2
9	31	2.0	538	41	AI584760	AI584760 fb83a09.y
10	31	2.0	545	79	AM649105	AM649105 EST327559
11	31	2.0	600	79	AM624043	AM624043 EST321988
12	31	2.0	628	79	AM649317	AM649317 EST327771
13	31	2.0	650	80	C24646	C24646 C24646 Dict
14	31	1.9	113	101	AO349263	AO349263 RPT11-13
15	30	1.9	198	123	HSU21500	U21500 Human Chrom
16	30	1.9	240	48	AV074622	AV074622 AV074622
17	30	1.9	267	117	AO927668	AO927668 RPT-23-2
18	30	1.9	323	120	BI7381	BI7381 345H6.TVB.C
19	30	1.9	345	94	AO067012	AO067012 HS_2236.A
20	30	1.9	344	94	AO106671	AO106671 HS_3080.B
21	30	1.9	349	118	AO988787	AO988787 OB1C03NE
22	30	1.9	359	88	R23641	R23641 YH35602.R1
23	30	1.9	363	36	AI174213	AI174213 an01c05.x
24	30	1.9	366	95	AO182320	AO182320 HS_2240.A
25	30	1.9	370	22	AA176610	AA176610 zp10e06.s
26	30	1.9	388	93	AO024546	AO024546 HS_2182.A
27	30	1.9	389	79	AM676174	AM676174 832008C08
28	30	1.9	393	85	F18858	F18858 HSPD03455.H
29	30	1.9	397	105	AO673094	AO673094 HS_5333.B
30	30	1.9	399	94	AO118216	AO118216 HS_3003.A
31	30	1.9	403	47	AL121094	AL121094 DKF2P762D
32	30	1.9	404	28	AA611235	AA611235 vna4911.r
33	30	1.9	407	86	H85715	H85715 Y668G06.r1
34	30	1.9	414	122	FR0022276	AL015146 F.rubripes
35	30	1.9	416	86	W78296	W78296 EST00444.Fe
36	30	1.9	449	34	AI059416	AI059416 UT-R-Cl-X
37	30	1.9	462	105	AO671317	AO671317 HS_5452.A
38	30	1.9	498	70	AM292326	AM292326 UT-H-B12-
39	30	1.9	509	70	AM301250	AM301250 X569E08.x
40	30	1.9	510	63	AM051367	AM051367 ME000327.
41	30	1.9	514	92	Z25512	Z25512 ARTS1025.GT
42	30	1.9	520	44	AI822995	AI822995 L30-846T3
43	30	1.9	521	89	R76963	R76963 Y166G07.r1
44	30	1.9	530	102	AO406707	AO406707 HS_5091.A
45	30	1.9	531	46	AI935442	AI935442 w084e06.x
46	30	1.9	537	103	AO516617	AO516617 HS_5081.B
47	30	1.9	539	80	C23689	C23689 C23689 Dict
48	30	1.9	553	113	AO699969	AO699969 HS_5334.A
49	30	1.9	553	122	FR0012436	AL003688 F.rubripes
50	30	1.9	563	101	AO386223	AO386223 RPT11-14
51	30	1.9	587	101	AO328711	AO328711 nxb0043P
52	30	1.9	589	47	AL049105	AL049105 DKF2P434C
53	30	1.9	594	105	AO635376	AO635376 Sheared D
54	30	1.9	596	105	AO639317	AO639317 927P1-1F1
55	30	1.9	597	31	AA836741	AA836741 L30-189T3

56	30	1.9	620	44	AI822151	AI822151 L0-626T3
57	30	1.9	626	45	AI861001	AI861001 L30-121ST
58	30	1.9	632	34	AI058056	AI058056 SWAMCA12
59	30	1.9	638	47	AD033322	AD033322 AD033322
60	30	1.9	648	101	AO379656	AO379656 RPT11-16
61	30	1.9	680	44	AI822243	AI822243 L0-619T3
62	30	1.9	687	103	AO481998	AO481998 RPT-11-2
63	30	1.9	718	101	AO349265	AO349265 RPT11-13
64	30	1.9	720	122	CNS007S	AL095830 Arabidops
65	30	1.9	747	96	AO271265	AO271265 nxb0025B
66	30	1.9	759	113	AO738655	AO738655 HS_3581.B
67	30	1.9	777	122	AG010424	AG010424 Homo sap1
68	30	1.9	789	122	AG010425	AG010425 Homo sap1
69	30	1.9	798	63	AM054390	AM054390 L30-1728T
70	30	1.9	808	122	AG010423	AG010423 Homo sap1
71	30	1.9	1101	122	CNS00PQC	AL070489 Drosophill
72	30	1.9	1101	123	CNS0151G	AL104878 Drosophill
73	29	1.9	104	96	AO323646	AO323646 RPT11-93
74	29	1.9	159	119	AZ074059	AZ074059 RPT-23-4
75	29	1.9	170	47	AL038873	AL038873 DKF2P566C
76	29	1.9	171	119	AZ060520	AZ060520 RPT-23-4
77	29	1.9	181	120	B47962	B47962 RPT11-3M6.
78	29	1.9	183	69	AM232081	AM232081 f13d08.x
79	29	1.9	210	103	AO504031	AO504031 RPT-11-2
80	29	1.9	213	48	AO039832	AO039832 AU039832
81	29	1.9	214	73	AM522927	AM522927 UT-R-B00-
82	29	1.9	215	48	AO606841	AO606841 AU060841
83	29	1.9	240	48	AD073015	AD073015 AU073015
84	29	1.9	257	120	B47934	B47934 RPT11-3J5.
85	29	1.9	264	22	AA161725	AA161725 MNAFCR8G0
86	29	1.9	267	38	AI370356	AI370356 GMA5908.X
87	29	1.9	291	101	AO392457	AO392457 CTBT-ET-1
88	29	1.9	292	26	AA16327	AA16327 v01e08.s
89	29	1.9	293	42	AI647245	AI647245 v04911.Y
90	29	1.9	301	46	AI941621	AI941621 F18032C05
91	29	1.9	301	122	FR0010719	AL004073 F.rubripes
92	29	1.9	302	62	AV388558	AV388558 AV388558
93	29	1.9	310	102	AO439496	AO439496 HS_5091.A
94	29	1.9	312	122	CNS00PQM	AL084880 Arabidops
95	29	1.9	317	29	AA629963	AA629963 v760B08.S
96	29	1.9	318	103	AO471898	AO471898 CTBT-ET-1
97	29	1.9	325	62	AM018981	AM018981 f047H04.x
98	29	1.9	329	96	AO267717	AO267717 RPT11-76
99	29	1.9	345	91	W78689	W78689 EST00008.Tf
100	29	1.9	351	120	B43953	B43953 HS-1058-B1-
101	29	1.9	351	120	B24410	B24410 F20A16TR.IG
102	29	1.9	366	69	AM238020	AM238020 xp15d12.x
103	29	1.9	368	22	AA190731	AA190731 ZP88B08.r
104	29	1.9	377	90	T95763	T95763 Y640E10.s1
105	29	1.9	383	101	AO384343	AO384343 RPT11-14
106	29	1.9	388	122	FR0012823	AL004073 F.rubripes
107	29	1.9	395	120	B72162	B72162 RPT11-7J16
108	29	1.9	407	93	AO078535	AO078535 CIR-HSP-2
109	29	1.9	408	42	AI626774	AI626774 f03C01.Y
110	29	1.9	410	103	AO485566	AO485566 RPT-11-2
111	29	1.9	413	42	AI641313	AI641313 f034005.Y
112	29	1.9	413	119	AZ076880	AZ076880 RPT-23-4
113	29	1.9	419	42	AI641078	AI641078 f08903.x
114	29	1.9	421	95	AO211715	AO211715 HS_3241.B
115	29	1.9	421	120	B23010	B23010 F16H4P.IGF
116	29	1.9	422	34	AI014136	AI014136 EST207691
117	29	1.9	423	39	AI454014	AI454014 UT-R-BT0-
118	29	1.9	430	89	R95529	R95529 SMNHADA0180
119	29	1.9	436	73	AM521931	AM521931 UT-R-B00-
120	29	1.9	437	122	FR0006407	Z90217 F.rubripes
121	29	1.9	440	87	AA43111	AA43111 SW3ICA946SK
122	29	1.9	443	27	AA497294	AA497294 fa03e10.r
123	29	1.9	445	117	AO877592	AO877592 HS_2146.B
124	29	1.9	447	64	AM164001	AM164001 L11ndest
125	29	1.9	447	71	AW332785	AW332785 S13A4.AGS
126	29	1.9	449	120	B41863	B41863 HS-1054-B1-
127	29	1.9	452	45	AI871451	AI871451 w171h02.x
128	29	1.9	456	31	AA850442	AA850442 EST193209

C 129	29	1.9	457	33	AA933418	SWEML3SBH	202	29	1.9	688	81	C93984	C93984	DICT
C 130	29	1.9	457	96	AA0270966	HS-2059_A	203	29	1.9	697	64	AA134038	AA134038	F.L14G03.Y
C 131	29	1.9	459	45	AA1870349	W12H02.X	204	29	1.9	697	119	AA069987	AA069987	RPCI-23-4
C 132	29	1.9	461	31	AA819896	UT-R-A0-a	205	29	1.9	714	79	AA622250	AA622250	EST313048
C 133	29	1.9	461	43	AA1706038	UT-R-AB1-	206	29	1.9	720	103	AA085905	AA085905	RPCI-11-2
C 134	29	1.9	464	114	AA0870506	HS-3248_A	207	29	1.9	721	48	AA050420	AA050420	AA050420
C 135	29	1.9	466	36	AA1217303	q125d01.X	208	29	1.9	731	104	AA0576427	AA0576427	nbx50089K
C 136	29	1.9	470	40	AA1013894	EST208569	209	29	1.9	733	118	AA005271	AA005271	AA005271
C 137	29	1.9	474	34	AA1507433	v134d11.X	210	29	1.9	733	118	AA007296	AA007296	RPCI-23-3
C 138	29	1.9	475	73	AA0525209	UT-R-B10P	211	29	1.9	738	101	AA083523	AA083523	RPCI-11-13
C 139	29	1.9	482	37	AA1239481	EST225176	212	29	1.9	776	47	AA000204	AA000204	AA000204
C 140	29	1.9	483	27	AA0542455	EST07B10.S	213	29	1.9	783	34	AA1069169	AA1069169	AA1069169
C 141	29	1.9	488	34	AA1044654	UT-R-C1-1	214	29	1.9	814	47	AA005250	AA005250	AA005250
C 142	29	1.9	493	41	AA1568452	tn39d10.X	215	29	1.9	814	122	AA010456	AA010456	Homc sap1
C 143	29	1.9	496	41	AA1539334	ts45q09.X	216	29	1.9	1008	46	AA1943438	AA1943438	MP02E10.M
C 144	29	1.9	498	101	AA0359777	HS-5027_A	217	29	1.9	1101	122	CNS00067	CNS00067	AL050601.Drosophila
C 145	29	1.9	502	69	AA174086	q139d03.Y	218	29	1.9	1101	122	CNS000DP	CNS000DP	AL050601.Drosophila
C 146	29	1.9	504	69	AA0233209	AA0233209	219	29	1.9	1137	120	B10133	B10133	F2H22-T7 IG
C 147	29	1.9	504	101	AA0387086	RPCI11-14	220	29	1.9	109	23	AA247799	AA247799	10406.seq
C 148	29	1.9	506	101	AA012734	AA012734	221	29	1.8	190	47	AA036634	AA036634	DKE25564A
C 149	29	1.9	506	101	AA0339181	HS-5022_A	222	29	1.8	241	44	AA1924565	AA1924565	AA1924565
C 150	29	1.9	505	119	AA027677	RPCI-23-3	223	29	1.8	252	101	AA0353022	AA0353022	AA0353022
C 151	29	1.9	509	35	AA1111796	UT-R-10-m	224	29	1.8	268	80	C27777	C27777	AA1866380.tu41C04.X
C 152	29	1.9	512	74	AA0359402	AA0359402	225	29	1.8	308	79	AA0684341	AA0684341	AA0684341
C 153	29	1.9	512	105	AA067130	HS-2106_B	226	29	1.8	318	28	AA0587886	AA0587886	AA0587886
C 154	29	1.9	522	80	C25513	C25513	227	29	1.8	326	48	AA061116	AA061116	AA061116
C 155	29	1.9	522	123	AA0450943	AA0450943	228	29	1.8	331	43	AA1686380	AA1686380	AA1686380
C 156	29	1.9	525	102	AA0450943	AA0450943	229	29	1.8	351	24	AA0297731	AA0297731	AA0297731
C 157	29	1.9	529	36	AA1168932	AA1168932	230	29	1.8	351	24	AA0297731	AA0297731	AA0297731
C 158	29	1.9	529	105	AA015950	RPCI-23-3	231	29	1.8	360	81	D34916	D34916	AA1967598
C 159	29	1.9	530	119	AA0639023	927P1-10C	232	29	1.8	361	38	AA1372496	AA1372496	AA1372496
C 160	29	1.9	531	41	AA1588295	AA1588295	233	29	1.8	381	38	AA096138	AA096138	AA096138
C 161	29	1.9	531	72	AA0462445	BP230009A	234	29	1.8	388	94	AA096138	AA096138	AA096138
C 162	29	1.9	531	103	AA0484742	RPCI-11-2	235	29	1.8	398	32	AA026136	AA026136	AA026136
C 163	29	1.9	532	79	AA0637548	B159903.W	236	29	1.8	406	40	AA1507995	AA1507995	AA1507995
C 164	29	1.9	537	101	AA0357863	CITB1-EI-	237	29	1.8	409	73	AA0532460	AA0532460	AA0532460
C 165	29	1.9	538	101	AA0392654	CITB1-EI-	238	29	1.8	413	37	AA1242126	AA1242126	AA1242126
C 166	29	1.9	539	103	AA0468796	HS-5117_A	239	29	1.8	418	43	AA1706362	AA1706362	AA1706362
C 167	29	1.9	547	120	B63386	B63386	240	29	1.8	428	118	AA0959206	AA0959206	AA0959206
C 168	29	1.9	548	120	AA0809997	RPCI-23-4	241	29	1.8	440	81	C95421	C95421	C95421
C 169	29	1.9	551	43	AA1729897	BMDG1530	242	29	1.8	444	39	AA1397898	AA1397898	AA1397898
C 170	29	1.9	561	117	AA0919456	RPCI-23-2	243	29	1.8	449	102	AA0465958	AA0465958	AA0465958
C 171	29	1.9	563	119	AA016768	RPCI-23-2	244	29	1.8	463	28	AA0606417	AA0606417	AA0606417
C 172	29	1.9	563	120	AA080499	RPCI-23-4	245	29	1.8	463	95	AA0250014	AA0250014	AA0250014
C 173	29	1.9	566	117	AA0881369	HS-5158_A	246	29	1.8	472	116	AA0820134	AA0820134	AA0820134
C 174	29	1.9	568	43	AA1730165	BMDG1633	247	29	1.8	481	95	AA0221942	AA0221942	AA0221942
C 175	29	1.9	571	48	AA0061968	AA0061968	248	29	1.8	484	105	AA0188662	AA0188662	AA0188662
C 176	29	1.9	574	48	AA0061968	AA0061968	249	29	1.8	498	41	AA1574537	AA1574537	AA1574537
C 177	29	1.9	575	103	AA0522898	AA0522898	250	29	1.8	499	122	AA0008710	AA0008710	AA0008710
C 178	29	1.9	575	122	AA0012622	F.rubripes	251	29	1.8	507	37	AA1283987	AA1283987	AA1283987
C 179	29	1.9	577	123	AA0012622	F.rubripes	252	29	1.8	507	70	AA0254631	AA0254631	AA0254631
C 180	29	1.9	580	123	AA0036775	AA0036775	253	29	1.8	511	63	AA0808809	AA0808809	AA0808809
C 181	29	1.9	586	122	AA0036775	AA0036775	254	29	1.8	513	72	AA0435347	AA0435347	AA0435347
C 182	29	1.9	591	46	AA1977595	AA1977595	255	29	1.8	519	37	AA1237205	AA1237205	AA1237205
C 183	29	1.9	599	71	AA0333565	AA0333565	256	29	1.8	536	70	AA0532658	AA0532658	AA0532658
C 184	29	1.9	608	122	AA0031337	AA0031337	257	29	1.8	540	39	AA1398762	AA1398762	AA1398762
C 185	29	1.9	611	102	AA0397675	AA0397675	258	29	1.8	551	28	AA018901	AA018901	AA018901
C 186	29	1.9	611	118	AA061503	AA061503	259	29	1.8	560	39	AA1398663	AA1398663	AA1398663
C 187	29	1.9	617	43	AA1727264	AA1727264	260	29	1.8	565	118	AA059204	AA059204	AA059204
C 188	29	1.9	618	117	AA0879649	AA0879649	261	29	1.8	570	39	AA1398111	AA1398111	AA1398111
C 189	29	1.9	619	103	AA0530235	RPCI-11-3	262	29	1.8	573	96	AA0256565	AA0256565	AA0256565
C 190	29	1.9	621	71	AA0335716	AA0335716	263	29	1.8	574	101	AA0385797	AA0385797	AA0385797
C 191	29	1.9	632	45	AA1877859	AA1877859	264	29	1.8	574	118	AA059207	AA059207	AA059207
C 192	29	1.9	632	45	AA1877859	AA1877859	265	29	1.8	583	101	AA0390526	AA0390526	AA0390526
C 193	29	1.9	633	101	AA0375166	AA0375166	266	29	1.8	584	44	AA1765480	AA1765480	AA1765480
C 194	29	1.9	636	43	AA1726233	AA1726233	267	29	1.8	597	63	AA0449311	AA0449311	AA0449311
C 195	29	1.9	639	43	AA1731039	AA1731039	268	29	1.8	614	103	AA0491350	AA0491350	AA0491350
C 196	29	1.9	650	118	AA0955584	AA0955584	269	29	1.8	615	80	C25756	C25756	C25756
C 197	29	1.9	650	118	AA0955584	AA0955584	270	29	1.8	619	122	AA06904496	AA06904496	AA06904496
C 198	29	1.9	656	101	AA0386376	RPCI11-15	271	29	1.8	623	113	AA0690132	AA0690132	AA0690132
C 199	29	1.9	676	101	AA0352186	CITB1-EI-	272	29	1.8	634	113	AA0696421	AA0696421	AA0696421
C 200	29	1.9	679	79	AA0689203	AA0689203	273	29	1.8	636	47	AA029338	AA029338	AA029338
C 201	29	1.9	681	121	AA08606	AA08606	274	29	1.8	637	48	AA060724	AA060724	AA060724
C 202	29	1.9	686	47	AA036244	DKE2P564K	275	29	1.8	637	48	AA060724	AA060724	AA060724

275	28	1.8	638	72	AM460744	da35d01.x	c 348	26	1.7	348	25	AA391426	AA391426 LD09857.5
276	28	1.8	642	101	AQ327713	nbxb0041L	c 349	26	1.7	349	85	H04669	H04669 y149f05.r1
277	28	1.8	645	96	AQ258181	nbxb0019H	c 350	26	1.7	350	103	AQ473193	AQ473193 CIRB1-E1
278	28	1.8	665	47	AL037005	DKF2P564G	c 351	26	1.7	400	38	AT353515	AT353515 zeh0619.s
279	28	1.8	680	48	AU061831	AU061831	c 352	26	1.7	404	34	AT061530	AT061530 SMOVL3CAN
280	28	1.8	688	47	AU033562	AU033562	c 353	26	1.7	411	26	AA438964	AA438964 LD13368.5
281	28	1.8	701	48	AU060702	AU060702	c 354	26	1.7	415	102	AQ438967	AQ438967 HS_-5051.B
282	28	1.8	714	102	AQ438389	HS_-5109_B	c 355	26	1.7	415	105	AQ657114	AQ657114 sheared D
283	28	1.8	732	118	AQ959205	AQ959205 LERH82TF	c 356	26	1.7	450	103	AQ488524	AQ488524 RPT-11-2
284	28	1.8	848	120	B21148	B21148 F20B13-T7 I	c 357	26	1.7	458	69	AM185167	AM185167 se87g05.y
285	27	1.7	184	123	FR0041316	FR0041316	c 358	26	1.7	461	45	AT899958	AT899958 sb96g11.y
286	27	1.7	193	42	AT641515	AT641515 fc16f10.x	c 359	26	1.7	466	93	AQ010505	AQ010505 F23B3TEC
287	27	1.7	231	87	MA9326	MA9326 yv45g10.r1	c 360	26	1.7	482	71	AA328821	AA328821 N200012e
288	27	1.7	229	31	AA829568	od97D03.s	c 361	26	1.7	509	103	AQ488285	AQ488285 RPT-11-2
289	27	1.7	262	80	AA698894	AA698894 r167 non-	c 362	26	1.7	513	104	AQ595169	AQ595169 HS_5424.B
290	27	1.7	265	25	AA373519	AA373519 ESTP5679	c 363	26	1.7	518	94	AQ151115	AQ151115 HS_2185_B
291	27	1.7	343	101	AQ365564	AQ365564 nbxb0064G	c 364	26	1.7	518	103	AQ489602	AQ489602 RPT-11-2
292	27	1.7	345	69	AM32242	AM32242 f116d05.x	c 365	26	1.7	519	35	AT122848	AT122848 qb01b05.s
293	27	1.7	366	69	AM194050	AM194050 xm11d10.x	c 366	26	1.7	519	37	AT129314	AT129314 GH16341.3
294	27	1.7	370	123	AF046247	AF046247 Mus muscu	c 367	26	1.7	520	41	AT584979	AT584979 fb93f11.y
295	27	1.7	379	40	AT500821	AT500821 UI-R-C2B-	c 368	26	1.7	520	87	MA7911	MA7911 yw97a12.r1
296	27	1.7	382	85	H37415	H37415 15544.Lambd	c 369	26	1.7	530	35	AT100617	AT100617 34792.Lam
297	27	1.7	388	38	AT352188	AT352188 qrt10h08.x	c 370	26	1.7	531	113	AQ683748	AQ683748 HS_-3455_B
298	27	1.7	390	31	AA848939	AA848939 EST191701	c 371	26	1.7	540	105	AQ611316	AQ611316 HS_-5088_B
299	27	1.7	409	95	AQ231950	HS_3243_B	c 372	26	1.7	546	26	AA397479	AA397479 CPEST_269
300	27	1.7	419	95	AA170294	AA170294 EST216220	c 373	26	1.7	555	96	AQ320945	AQ320945 RPT-11-10
301	27	1.7	434	21	AA083088	AA083088 zn08h06.s	c 374	26	1.7	558	37	AT1239179	AT1239179 GM05128.5
302	27	1.7	440	36	AT182515	AT182515 ue21g12.x	c 375	26	1.7	564	43	AT7227363	AT7227363 BNGH1784
303	27	1.7	442	20	AA053415	AA053415 z171d03.s	c 376	26	1.7	580	101	AQ368741	AQ368741 RPT-11-14
304	27	1.7	463	95	AQ178867	AQ178867 HS_-3061.B	c 377	26	1.7	580	123	FR0042113	FR0042113 Fugu rubr
305	27	1.7	466	27	AA524108	AA524108 ng35a11.s	c 378	26	1.7	599	123	FR0042116	FR0042116 Fugu rubr
306	27	1.7	469	71	AM341789	AM341789 ha02b07.x	c 379	26	1.7	601	123	FR0042118	FR0042118 Fugu rubr
307	27	1.7	483	40	AT508023	AT508023 BSBML3520	c 380	26	1.7	602	123	FR0042309	FR0042309 Fugu rubr
308	27	1.7	501	102	AQ409888	AQ409888 HS_5096_B	c 381	26	1.7	604	62	AM017299	AM017299 EST72721
309	27	1.7	505	30	AA699207	AA699207 HU07930.3	c 382	26	1.7	616	48	AU050060	AU050060 AU050060
310	27	1.7	517	81	C92940	C92940 C92940 D1CT	c 383	26	1.7	618	104	AQ592123	AQ592123 HS_-5418_B
311	27	1.7	523	105	AQ654998	AQ654998 Sheared D	c 384	26	1.7	626	48	AU050081	AU050081 AU050081
312	27	1.7	524	23	AA224679	AA224679 CPEST_146	c 385	26	1.7	632	29	AA698618	AA698618 HL05343.5
313	27	1.7	524	26	AA427938	AA427938 zw53d03.s	c 386	26	1.7	638	74	AA585158	AA585158 EST317781
314	27	1.7	524	41	AT584171	AT584171 ts14f07.x	c 387	26	1.7	643	118	AQ950748	AQ950748 sheared D
315	27	1.7	525	42	AT617736	AT617736 zehm2050.	c 388	26	1.7	654	79	AM687036	AM687036 NF005C07R
316	27	1.7	528	101	AQ379761	AQ379761 RPT-11-15	c 389	26	1.7	661	118	AQ696384	AQ696384 LERJW3TF
317	27	1.7	565	28	AA610066	AA610066 AT08e07.s	c 390	26	1.7	661	118	AQ696385	AQ696385 LERJW3TR
318	27	1.7	573	43	AT742063	AT742063 wg38e06.x	c 391	26	1.7	665	79	AM685901	AM685901 NF031E11
319	27	1.7	589	91	W56087	W56087 zc66f08.r1	c 392	26	1.7	693	93	AQ075642	AQ075642 CIT-HS-2
320	27	1.7	605	69	AM186938	AM186938 BNLGH1753	c 393	26	1.7	707	117	AQ914179	AQ914179 nbeD0047F
321	27	1.7	610	39	AT397825	AT397825 NCM7A917	c 394	26	1.7	733	101	AQ325736	AQ325736 nbxb0020E
322	27	1.7	615	118	AQ961548	AQ961548 LERFL89TF	c 395	26	1.7	751	36	AT1188382	AT1188382 qd13f08.x
323	27	1.7	620	81	C83832	C83832 C83832 D1CT	c 396	26	1.7	769	23	AA263375	AA263375 LD06365.5
324	27	1.7	624	48	AU039020	AU039020 AU039020	c 397	26	1.7	784	48	AU050701	AU050701 AU050701
325	27	1.7	630	81	C90182	C90182 C90182 D1CT	c 398	26	1.7	802	48	AU050688	AU050688 AU050688
326	27	1.7	635	42	AT626367	AT626367 fc04g12.x	c 399	26	1.7	806	120	B05687	B05687
327	27	1.7	636	62	AV439514	AV439514 AV439514	c 400	26	1.7	818	26	AA439858	AA439858
328	27	1.7	637	102	AQ450439	AQ450439 500011G06	c 401	26	1.7	824	62	AV384393	AV384393
329	27	1.7	641	118	AQ966734	AQ966734 LERL183TF	c 402	26	1.7	840	81	C82120	C82120 Leuk
330	27	1.7	655	44	AT793516	AT793516 fc50e01.x	c 403	26	1.7	860	81	C82184	C82184 C82184 Leuk
331	27	1.7	659	81	C92104	C92104 C92104 D1CT	c 404	26	1.7	899	122	CNS008C8	CNS008C8
332	27	1.7	667	64	AM154541	AM154541 f130e11.x	c 405	26	1.7	1107	80	C22922	C22922 D1CT
333	27	1.7	689	118	AQ961549	AQ961549 LERFL89TF	c 406	26	1.7	117	25	AA368278	AA368278 EST79506
334	27	1.7	737	118	AQ955486	AQ955486 LERAE30TF	c 407	26	1.7	125	32	AA879214	AA879214 nb85d06.s
335	27	1.7	740	117	AQ935668	AQ935668 CPg2588A	c 408	26	1.7	240	48	AU071681	AU071681 AU071681
336	27	1.7	744	118	AQ955487	AQ955487 LERAE30TR	c 409	26	1.7	240	48	AU072466	AU072466 AU072466
337	27	1.7	751	120	B19195	B19195 FSG12-T7 IG	c 410	26	1.7	274	31	AA782068	AA782068 a147b12.s
338	27	1.7	822	116	AQ854846	AQ854846 CPg2022B	c 411	26	1.7	279	69	AM233617	AM233617 f339d08.x
339	27	1.7	829	122	CNS009VS	CNS009VS Drosoph11	c 412	26	1.7	283	46	AT1914766	AT1914766 t102b04.x
340	27	1.7	920	117	AQ935049	AQ935049 CPg2305B	c 413	26	1.7	284	104	AQ584290	AQ584290 RPT-11-4
341	27	1.7	992	122	CNS00L5T	CNS00L5T Drosoph11	c 414	26	1.7	294	33	AA932668	AA932668 co074a02.s
342	27	1.7	1101	122	CNS00EPZ	CNS00EPZ Drosoph11	c 415	26	1.7	322	122	FR0034412	FR0034412 Fugu rubr
343	27	1.7	1101	122	CNS00EVP	CNS00EVP Drosoph11	c 416	26	1.7	325	118	AZ013544	AZ013544 RPT-23-2
344	27	1.7	1101	123	CNS0160W	CNS0160W Drosoph11	c 417	26	1.7	365	24	AA296820	AA296820 EST112476
345	26	1.7	164	20	AA035437	AA035437 zk25f12.r	c 418	26	1.7	369	105	AQ658874	AQ658874 sheared D
346	26	1.7	265	48	AU061291	AU061291	c 419	26	1.7	380	93	AQ038570	AQ038570 CIT-HSP-2
347	26	1.7	323	38	AT366026	AT366026 a082d06.x	c 420	26	1.7	386	92	Z26004	Z26004 ATTS1281.GI

421	25	1.6	400	93	AO05010	494	24	1.5	306	89	T20901	T20901 2909 Lambda
422	25	1.6	401	117	AO908519	495	24	1.5	317	62	AV442612	AV442612
423	25	1.6	406	87	N97211	496	24	1.5	329	85	F20810	F20810 HSP05216 H
424	25	1.6	406	104	AO566120	497	24	1.5	330	117	AO908408	AO908408 GSTC0516
425	25	1.6	409	96	AO306679	498	24	1.5	336	40	AI502302	AI502302 UT-R-C1-1
426	25	1.6	412	96	T76076	499	24	1.5	337	48	AO061365	AO061365
427	25	1.6	416	86	H51991	500	24	1.5	339	48	AO062178	AO062178
428	25	1.6	417	117	AO884167	501	24	1.5	346	93	AO010969	AO010969 F23167R
429	25	1.6	418	93	AO023443	502	24	1.5	352	33	AA999250	AA999250 MCE3044
430	25	1.6	419	36	AI191724	503	24	1.5	358	29	AA648240	AA648240 nst9070.T
431	25	1.6	420	29	AA681031	504	24	1.5	382	122	FR0014460	FR0014460
432	25	1.6	431	79	AM636342	505	24	1.5	392	72	AW469011	AW469011 hc74f12.x
433	25	1.6	433	86	N22920	506	24	1.5	404	88	RI1655	RI1655 Y43C09.r1
434	25	1.6	448	87	N73559	507	24	1.5	417	94	AO144515	AO144515 HS_3078_B
435	25	1.6	462	81	C72849	508	24	1.5	418	87	N38075	N38075 19302 Lambda
436	25	1.6	466	47	AL135589	509	24	1.5	425	105	AO616436	AO616436 HS_5149_B
437	25	1.6	470	47	AL0336418	510	24	1.5	440	33	AI415924	AI415924 fb36h02.x
438	25	1.6	471	120	B74195	511	24	1.5	443	33	AA945672	AA945672 EST201171
439	25	1.6	479	22	AA167097	512	24	1.5	447	45	AI884085	AI884085 fc73d10.x
440	25	1.6	481	64	AM116435	513	24	1.5	455	46	AI965323	AI965323 fc89h04.x
441	25	1.6	508	47	AO034669	514	24	1.5	460	35	AI093407	AI093407 qd03906.x
442	25	1.6	510	81	C84881	515	24	1.5	460	35	AO040055	AO040055
443	25	1.6	514	81	AI878670	516	24	1.5	461	46	AI965057	AI965057 fc86c04.Y
444	25	1.6	530	105	AO613026	517	24	1.5	466	80	AW719789	AW719789 L1NEST10A
445	25	1.6	533	22	AA166771	518	24	1.5	471	48	AM051808	AM051808
446	25	1.6	544	71	AM370434	519	24	1.5	471	69	AM198054	AM198054 xc27h09.x
447	25	1.6	551	103	AO503079	520	24	1.5	476	96	AO235210	AO235210 HS_-2008_B
448	25	1.6	556	102	AO440742	521	24	1.5	479	34	AI026298	AI026298 LO-515T3
449	25	1.6	564	81	C84083	522	24	1.5	485	47	AO031455	AO031455
450	25	1.6	578	118	AO689907	523	24	1.5	498	103	AO489633	AO489633 RPCI-11-2
451	25	1.6	579	118	AO958817	524	24	1.5	505	101	AO341236	AO341236 RPCI-11-12
452	25	1.6	585	102	AO603054	525	24	1.5	510	122	FR0004797	FR0004797
453	25	1.6	588	122	FR0022200	526	24	1.5	516	43	AI722302	AI722302 F_rubripes
454	25	1.6	597	47	AU029747	527	24	1.5	524	44	AI793769	AI793769 fc32c06.x
455	25	1.6	598	64	AM115825	528	24	1.5	525	120	B28060	B28060 T28887R TA
456	25	1.6	599	105	AO652064	529	24	1.5	530	38	AI353541	AI353541 zeh05128
457	25	1.6	614	71	AM349271	530	24	1.5	531	62	AA479272	AA479272 zv21807.S
458	25	1.6	616	118	AO959064	531	24	1.5	531	62	AV442718	AV442718 AV442718
459	25	1.6	623	105	AO638394	532	24	1.5	531	113	AO726743	AO726743 HS_-3407_B
460	25	1.6	628	105	AO638394	533	24	1.5	540	40	AO034191	AO034191
461	25	1.6	650	71	AM370396	534	24	1.5	549	62	AM019333	AM019333 fc36e12.Y
462	25	1.6	650	118	AO958694	535	24	1.5	553	43	AI794035	AI794035
463	25	1.6	650	118	AO958694	536	24	1.5	557	93	AO024003	AO024003 CPG8042A
464	25	1.6	667	40	AI525056	537	24	1.5	568	42	AI626342	AI626342 fc3h03.x
465	25	1.6	668	91	W78683	538	24	1.5	576	48	AI943097	AI943097 fc84d04.x
466	25	1.6	669	118	AO966728	539	24	1.5	582	46	AO052810	AO052810
467	25	1.6	712	118	AO958695	540	24	1.5	592	32	AA887425	AA887425 c150b02.s
468	25	1.6	712	118	AO958695	541	24	1.5	594	42	AI674388	AI674388 wc44a01.x
469	25	1.6	751	34	AI055104	542	24	1.5	602	48	AO060156	AO060156
470	25	1.6	760	114	AO744944	543	24	1.5	604	48	AO061478	AO061478
471	25	1.6	832	122	CNS00AHQ	544	24	1.5	608	96	AO256358	AO256358 nbx0016M
472	25	1.6	877	122	CNS00AE7	545	24	1.5	620	101	AO350104	AO350104 RPCI-11-10
473	25	1.6	935	122	CNS00502	546	24	1.5	624	119	AO53124	AO53124 RPCI-23-4
474	25	1.6	966	122	CNS0177U	547	24	1.5	638	47	AU029537	AU029537
475	25	1.6	982	122	CNS005T1	548	24	1.5	661	74	AM555783	AM555783 10259802-
476	25	1.6	1002	122	CNS0013V	549	24	1.5	677	48	AU061834	AU061834
477	25	1.6	1033	123	CNS01672	550	24	1.5	708	104	AO540590	AO540590 RPCI-11-3
478	25	1.6	1081	123	CNS0174Q	551	24	1.5	732	47	AU034184	AU034184
479	25	1.6	1101	122	CNS0024K	552	24	1.5	794	120	B21003	B21003 T3K16-T7 TA
480	25	1.6	1101	122	CNS007AT	553	24	1.5	849	122	CNS0082J	CNS0082J
481	25	1.6	1101	122	CNS00BAK	554	24	1.5	859	42	AO295337	AO295337
482	25	1.6	1101	122	CNS00FX9	555	24	1.5	1101	123	CNS014UF	CNS014UF
483	25	1.6	1101	122	CNS00GAV	556	24	1.5	147	95	AO194142	AO194142
484	25	1.6	1101	122	CNS00HHC	557	24	1.5	172	47	AL038878	AL038878
485	25	1.6	1101	123	CNS017WK	558	24	1.5	191	119	AZ068566	AZ068566 RPCI-23-4
486	25	1.5	191	48	AU060508	559	24	1.5	196	120	B55155	B55155 CIR-HSP-386
487	24	1.5	194	48	AU061757	560	24	1.5	204	113	AO680442	AO680442 HS_2275.B
488	24	1.5	221	38	AI310604	561	24	1.5	238	74	AM560873	AM560873 EST315921
489	24	1.5	223	42	AI669572	562	24	1.5	240	48	AU074697	AU074697
490	24	1.5	240	42	AU074303	563	24	1.5	242	47	AL036759	AL036759 DKF2P564J
491	24	1.5	277	89	T18065	564	24	1.5	296	64	AM127323	AM127323 M110494 D
492	24	1.5	294	122	CNS00XIC	565	24	1.5	328	121	CNS00UWF	CNS00UWF
493	24	1.5	300	51	AV199845	566	24	1.5	349	47	AU033553	AU033553

```

567 1.5 352 71 AM330121 TENU4867
568 1.5 353 89 T13748
569 1.5 372 63 AM035615
570 1.5 372 63 AM036481
571 1.5 386 33 AA943326
572 1.5 381 95 AO187955
573 1.5 396 25 AA660085
574 1.5 400 90 U74175
575 1.5 420 113 AO715351
576 1.5 421 104 AO583696
577 1.5 430 48 AO662506
578 1.5 432 32 AA899899
579 1.5 436 93 AO082314
580 1.5 436 116 AO819228
581 1.5 438 70 AM257100
582 1.5 438 96 AO260462
583 1.5 439 43 AT723141
584 1.5 438 69 AM186491
585 1.5 432 120 B41347
586 1.5 454 73 AM520860
587 1.5 454 81 C84063
588 1.5 461 118 AO12154
589 1.5 463 47 AT919151
590 1.5 473 48 AO060884
591 1.5 475 41 AT545084
592 1.5 476 30 AA701960
593 1.5 477 24 AA311223
594 1.5 479 48 AO060698
595 1.5 480 39 AT438346
596 1.5 482 69 AM184580
597 1.5 485 95 AO226769
598 1.5 486 79 AM617338
599 1.5 486 119 AO16647
600 1.5 499 96 AO324070
601 1.5 500 29 AA680492
602 1.5 502 27 AA494966
603 1.5 504 41 AT562823
604 1.5 508 37 AT122859
605 1.5 510 36 AT126935
606 1.5 512 38 AT381101
607 1.5 513 113 AO681844
608 1.5 515 72 AA635419
609 1.5 520 37 AT1234426
610 1.5 522 43 AT722191
611 1.5 537 29 AA651397
612 1.5 538 85 H36274
613 1.5 539 63 AM031842
614 1.5 540 41 AT584772
615 1.5 549 116 AO835000
616 1.5 552 37 AT259038
617 1.5 559 117 AO931651
618 1.5 556 48 AO060265
619 1.5 562 28 AA557106
620 1.5 572 42 AT167030
621 1.5 573 74 AA656713
622 1.5 578 25 AA393574
623 1.5 578 119 AO231903
624 1.5 579 34 AT0621903
625 1.5 581 47 AT280882
626 1.5 585 48 AO060740
627 1.5 586 101 AO394924
628 1.5 592 69 AM184579
629 1.5 596 37 AT292851
630 1.5 597 35 AT134828
631 1.5 601 70 AM306356
632 1.5 602 72 AA422391
633 1.5 610 64 AA143153
634 1.5 610 105 AO656209
635 1.5 611 35 AT111238
636 1.5 612 113 AO679182
637 1.5 617 48 AO037864
638 1.5 619 122 FRO008530
639 1.5 621 35 AT103950
640 1.5 640 47 AM330121 TENU4867
641 1.5 641 89 T13748
642 1.5 642 63 AM035615
643 1.5 642 63 AM036481
644 1.5 644 33 AA943326
645 1.5 645 95 AO187955
646 1.5 646 25 AA660085
647 1.5 647 90 U74175
648 1.5 647 113 AO715351
649 1.5 648 104 AO583696
650 1.5 649 48 AO662506
651 1.5 652 32 AA899899
652 1.5 652 93 AO082314
653 1.5 652 116 AO819228
654 1.5 652 70 AM257100
655 1.5 652 96 AO260462
656 1.5 652 43 AT723141
657 1.5 652 69 AM186491
658 1.5 652 120 B41347
659 1.5 652 73 AM520860
660 1.5 652 81 C84063
661 1.5 652 118 AO12154
662 1.5 652 47 AT919151
663 1.5 652 48 AO060884
664 1.5 652 41 AT545084
665 1.5 652 30 AA701960
666 1.5 652 24 AA311223
667 1.5 652 48 AO060698
668 1.5 652 39 AT438346
669 1.5 652 69 AM184580
670 1.5 652 95 AO226769
671 1.5 652 79 AM617338
672 1.5 652 119 AO16647
673 1.5 652 96 AO324070
674 1.5 652 29 AA680492
675 1.5 652 27 AA494966
676 1.5 652 41 AT562823
677 1.5 652 37 AT122859
678 1.5 652 36 AT126935
679 1.5 652 38 AT381101
680 1.5 652 113 AO681844
681 1.5 652 72 AA635419
682 1.5 652 37 AT1234426
683 1.5 652 43 AT722191
684 1.5 652 29 AA651397
685 1.5 652 85 H36274
686 1.5 652 63 AM031842
687 1.5 652 41 AT584772
688 1.5 652 116 AO835000
689 1.5 652 37 AT259038
690 1.5 652 117 AO931651
691 1.5 652 48 AO060265
692 1.5 652 28 AA557106
693 1.5 652 42 AT167030
694 1.5 652 74 AA656713
695 1.5 652 25 AA393574
696 1.5 652 119 AO231903
697 1.5 652 34 AT0621903
698 1.5 652 47 AT280882
699 1.5 652 48 AO060740
700 1.5 652 101 AO394924
701 1.5 652 69 AM184579
702 1.5 652 37 AT292851
703 1.5 652 35 AT134828
704 1.5 652 70 AM306356
705 1.5 652 72 AA422391
706 1.5 652 105 AO656209
707 1.5 652 113 AO679182
708 1.5 652 48 AO037864
709 1.5 652 122 FRO008530
710 1.5 652 35 AT103950
711 1.5 652 35 AT103950
712 1.5 652 35 AT103950
713 1.5 652 35 AT103950
714 1.5 652 35 AT103950
715 1.5 652 35 AT103950
716 1.5 652 35 AT103950
717 1.5 652 35 AT103950
718 1.5 652 35 AT103950
719 1.5 652 35 AT103950
720 1.5 652 35 AT103950
721 1.5 652 35 AT103950
722 1.5 652 35 AT103950
723 1.5 652 35 AT103950
724 1.5 652 35 AT103950
725 1.5 652 35 AT103950
726 1.5 652 35 AT103950
727 1.5 652 35 AT103950
728 1.5 652 35 AT103950
729 1.5 652 35 AT103950
730 1.5 652 35 AT103950
731 1.5 652 35 AT103950
732 1.5 652 35 AT103950
733 1.5 652 35 AT103950
734 1.5 652 35 AT103950
735 1.5 652 35 AT103950
736 1.5 652 35 AT103950
737 1.5 652 35 AT103950
738 1.5 652 35 AT103950
739 1.5 652 35 AT103950
740 1.5 652 35 AT103950
741 1.5 652 35 AT103950
742 1.5 652 35 AT103950
743 1.5 652 35 AT103950
744 1.5 652 35 AT103950
745 1.5 652 35 AT103950
746 1.5 652 35 AT103950
747 1.5 652 35 AT103950
748 1.5 652 35 AT103950
749 1.5 652 35 AT103950
750 1.5 652 35 AT103950
751 1.5 652 35 AT103950
752 1.5 652 35 AT103950
753 1.5 652 35 AT103950
754 1.5 652 35 AT103950
755 1.5 652 35 AT103950
756 1.5 652 35 AT103950
757 1.5 652 35 AT103950
758 1.5 652 35 AT103950
759 1.5 652 35 AT103950
760 1.5 652 35 AT103950
761 1.5 652 35 AT103950
762 1.5 652 35 AT103950
763 1.5 652 35 AT103950
764 1.5 652 35 AT103950
765 1.5 652 35 AT103950
766 1.5 652 35 AT103950
767 1.5 652 35 AT103950
768 1.5 652 35 AT103950
769 1.5 652 35 AT103950
770 1.5 652 35 AT103950
771 1.5 652 35 AT103950
772 1.5 652 35 AT103950
773 1.5 652 35 AT103950
774 1.5 652 35 AT103950
775 1.5 652 35 AT103950
776 1.5 652 35 AT103950
777 1.5 652 35 AT103950
778 1.5 652 35 AT103950
779 1.5 652 35 AT103950
780 1.5 652 35 AT103950
781 1.5 652 35 AT103950
782 1.5 652 35 AT103950
783 1.5 652 35 AT103950
784 1.5 652 35 AT103950
785 1.5 652 35 AT103950
786 1.5 652 35 AT103950
787 1.5 652 35 AT103950
788 1.5 652 35 AT103950
789 1.5 652 35 AT103950
790 1.5 652 35 AT103950
791 1.5 652 35 AT103950
792 1.5 652 35 AT103950
793 1.5 652 35 AT103950
794 1.5 652 35 AT103950
795 1.5 652 35 AT103950
796 1.5 652 35 AT103950
797 1.5 652 35 AT103950
798 1.5 652 35 AT103950
799 1.5 652 35 AT103950
800 1.5 652 35 AT103950
801 1.5 652 35 AT103950
802 1.5 652 35 AT103950
803 1.5 652 35 AT103950
804 1.5 652 35 AT103950
805 1.5 652 35 AT103950
806 1.5 652 35 AT103950
807 1.5 652 35 AT103950
808 1.5 652 35 AT103950
809 1.5 652 35 AT103950
810 1.5 652 35 AT103950
811 1.5 652 35 AT103950
812 1.5 652 35 AT103950
813 1.5 652 35 AT103950
814 1.5 652 35 AT103950
815 1.5 652 35 AT103950
816 1.5 652 35 AT103950
817 1.5 652 35 AT103950
818 1.5 652 35 AT103950
819 1.5 652 35 AT103950
820 1.5 652 35 AT103950
821 1.5 652 35 AT103950
822 1.5 652 35 AT103950
823 1.5 652 35 AT103950
824 1.5 652 35 AT103950
825 1.5 652 35 AT103950
826 1.5 652 35 AT103950
827 1.5 652 35 AT103950
828 1.5 652 35 AT103950
829 1.5 652 35 AT103950
830 1.5 652 35 AT103950
831 1.5 652 35 AT103950
832 1.5 652 35 AT103950
833 1.5 652 35 AT103950
834 1.5 652 35 AT103950
835 1.5 652 35 AT103950
836 1.5 652 35 AT103950
837 1.5 652 35 AT103950
838 1.5 652 35 AT103950
839 1.5 652 35 AT103950
840 1.5 652 35 AT103950
841 1.5 652 35 AT103950
842 1.5 652 35 AT103950
843 1.5 652 35 AT103950
844 1.5 652 35 AT103950
845 1.5 652 35 AT103950
846 1.5 652 35 AT103950
847 1.5 652 35 AT103950
848 1.5 652 35 AT103950
849 1.5 652 35 AT103950
850 1.5 652 35 AT103950
851 1.5 652 35 AT103950
852 1.5 652 35 AT103950
853 1.5 652 35 AT103950
854 1.5 652 35 AT103950
855 1.5 652 35 AT103950
856 1.5 652 35 AT103950
857 1.5 652 35 AT103950
858 1.5 652 35 AT103950
859 1.5 652 35 AT103950
860 1.5 652 35 AT103950
861 1.5 652 35 AT103950
862 1.5 652 35 AT103950
863 1.5 652 35 AT103950
864 1.5 652 35 AT103950
865 1.5 652 35 AT103950
866 1.5 652 35 AT103950
867 1.5 652 35 AT103950
868 1.5 652 35 AT103950
869 1.5 652 35 AT103950
870 1.5 652 35 AT103950
871 1.5 652 35 AT103950
872 1.5 652 35 AT103950
873 1.5 652 35 AT103950
874 1.5 652 35 AT103950
875 1.5 652 35 AT103950
876 1.5 652 35 AT103950
877 1.5 652 35 AT103950
878 1.5 652 35 AT103950
879 1.5 652 35 AT103950
880 1.5 652 35 AT103950
881 1.5 652 35 AT103950
882 1.5 652 35 AT103950
883 1.5 652 35 AT103950
884 1.5 652 35 AT103950
885 1.5 652 35 AT103950
886 1.5 652 35 AT103950
887 1.5 652 35 AT103950
888 1.5 652 35 AT103950
889 1.5 652 35 AT103950
890 1.5 652 35 AT103950
891 1.5 652 35 AT103950
892 1.5 652 35 AT103950
893 1.5 652 35 AT103950
894 1.5 652 35 AT103950
895 1.5 652 35 AT103950
896 1.5 652 35 AT103950
897 1.5 652 35 AT103950
898 1.5 652 35 AT103950
899 1.5 652 35 AT103950
900 1.5 652 35 AT103950
901 1.5 652 35 AT103950
902 1.5 652 35 AT103950
903 1.5 652 35 AT103950
904 1.5 652 35 AT103950
905 1.5 652 35 AT103950
906 1.5 652 35 AT103950
907 1.5 652 35 AT103950
908 1.5 652 35 AT103950
909 1.5 652 35 AT103950
910 1.5 652 35 AT103950
911 1.5 652 35 AT103950
912 1.5 652 35 AT103950
913 1.5 652 35 AT103950
914 1.5 652 35 AT103950
915 1.5 652 35 AT103950
916 1.5 652 35 AT103950
917 1.5 652 35 AT103950
918 1.5 652 35 AT103950
919 1.5 652 35 AT103950
920 1.5 652 35 AT103950
921 1.5 652 35 AT103950
922 1.5 652 35 AT103950
923 1.5 652 35 AT103950
924 1.5 652 35 AT103950
925 1.5 652 35 AT103950
926 1.5 652 35 AT103950
927 1.5 652 35 AT103950
928 1.5 652 35 AT103950
929 1.5 652 35 AT103950
930 1.5 652 35 AT103950
931 1.5 652 35 AT103950
932 1.5 652 35 AT103950
933 1.5 652 35 AT103950
934 1.5 652 35 AT103950
935 1.5 652 35 AT103950
936 1.5 652 35 AT103950
937 1.5 652 35 AT103950
938 1.5 652 35 AT103950
939 1.5 652 35 AT103950
940 1.5 652 35 AT103950
941 1.5 652 35 AT103950
942 1.5 652 35 AT103950
943 1.5 652 35 AT103950
944 1.5 652 35 AT103950
945 1.5 652 35 AT103950
946 1.5 652 35 AT103950
947 1.5 652 35 AT103950
948 1.5 652 35 AT103950
949 1.5 652 35 AT103950
950 1.5 652 35 AT103950
951 1.5 652 35 AT103950
952 1.5 652 35 AT103950
953 1.5 652 35 AT103950
954 1.5 652 35 AT103950
955 1.5 652 35 AT103950
956 1.5 652 35 AT103950
957 1.5 652 35 AT103950
958 1.5 652 35 AT103950
959 1.5 652 35 AT103950
960 1.5 652 35 AT103950
961 1.5 652 35 AT103950
962 1.5 652 35 AT103950
963 1.5 652 35 AT103950
964 1.5 652 35 AT103950
965 1.5 652 35 AT103950
966 1.5 652 35 AT103950
967 1.5 652 35 AT103950
968 1.5 652 35 AT103950
969 1.5 652 35 AT103950
970 1.5 652 35 AT103950
971 1.5 652 35 AT103950
972 1.5 652 35 AT103950
973 1.5 652 35 AT103950
974 1.5 652 35 AT103950
975 1.5 652 35 AT103950
976 1.5 652 35 AT103950
977 1.5 652 35 AT103950
978 1.5 652 35 AT103950
979 1.5 652 35 AT103950
980 1.5 652 35 AT103950
981 1.5 652 35 AT103950
982 1.5 652 35 AT103950
983 1.5 652 35 AT103950
984 1.5 652 35 AT103950
985 1.5 652 35 AT103950
986 1.5 652 35 AT103950
987 1.5 652 35 AT103950
988 1.5 652 35 AT103950
989 1.5 652 35 AT103950
990 1.5 652 35 AT103950
991 1.5 652 35 AT103950
992 1.5 652 35 AT103950
993 1.5 652 35 AT103950
994 1.5 652 35 AT103950
995 1.5 652 35 AT103950
996 1.5 652 35 AT103950
997 1.5 652 35 AT103950
998 1.5 652 35 AT103950
999 1.5 652 35 AT103950
1000 1.5 652 35 AT103950

```


713	22	1.4	356	28	AA597472	29743.Lam	786	22	1.4	464	45	A1900671
714	22	1.4	358	39	A1437800	SA39109.Y	787	22	1.4	466	105	AQ648121
715	22	1.4	360	39	A1440948	SA55508.Y	788	22	1.4	466	74	AA596313
716	22	1.4	361	39	A1441977	SA82810.Y	789	22	1.4	466	80	AW714087
717	22	1.4	363	35	A1099882	34035.Lam	790	22	1.4	467	70	AA307371
718	22	1.4	366	80	AW711925	AW711925.f	791	22	1.4	467	72	AA507534
719	22	1.4	367	39	A1460454	SA78903.Y	792	22	1.4	469	73	AA471716
720	22	1.4	369	40	A1495228	SP02605.Y	793	22	1.4	469	80	AA709421
721	22	1.4	370	40	A1476864	IP52805.Y	794	22	1.4	471	93	AQ018485
722	22	1.4	373	39	A1522921	SA92812.Y	795	22	1.4	475	74	AA595942
723	22	1.4	374	39	A1443245	SA46C03.Y	796	22	1.4	475	74	AA595942
724	22	1.4	375	39	A1441002	SA63F09.Y	797	22	1.4	475	122	AG020635
725	22	1.4	376	39	A1441959	SA83C08.Y	798	22	1.4	477	64	AA133045
726	22	1.4	376	39	A1496380	SP05M12.Y	799	22	1.4	477	70	AA257022
727	22	1.4	378	39	A1441419	SA59F10.Y	800	22	1.4	479	24	AA307165
728	22	1.4	378	86	WM1757	SMNHADA0170	801	22	1.4	479	46	A1939001
729	22	1.4	381	90	WM06763	SMNHADA0170	802	22	1.4	479	48	A1939001
730	22	1.4	382	81	D24300	RICR1688A	803	22	1.4	479	81	C90603
731	22	1.4	383	39	A1443189	SA45F02.Y	804	22	1.4	479	90	TS8511
732	22	1.4	385	104	AQ572155	HS-2108.B	805	22	1.4	481	74	AA595985
733	22	1.4	387	93	AQ078124	CIT-HSP-2	806	22	1.4	483	41	A1588172
734	22	1.4	390	70	AA264812	XQ35F04.X	807	22	1.4	483	105	AA642835
735	22	1.4	391	40	A1496675	SA196675.SP09M11.Y	808	22	1.4	489	32	AA901373
736	22	1.4	392	48	AU065387	AU065387	809	22	1.4	494	79	AA636870
737	22	1.4	393	63	AA099835	AA099835	810	22	1.4	497	96	AAQ292570
738	22	1.4	394	80	AW711094	SA90909.Y	811	22	1.4	500	40	AA187680
739	22	1.4	395	39	A1437782	SA39E06.Y	812	22	1.4	502	73	AA5082276
740	22	1.4	397	40	A1460770	SA69E409.Y	813	22	1.4	507	103	AAQ486497
741	22	1.4	399	80	AW710304	SA23E109.Y	814	22	1.4	510	46	A1974204
742	22	1.4	399	80	AW713441	SA80509.Y	815	22	1.4	511	64	AA133326
743	22	1.4	399	80	AW714575	SA1C12E.Y	816	22	1.4	511	104	AQ573201
744	22	1.4	399	80	AW714596	SA1E04E.Y	817	22	1.4	512	42	AA1658017
745	22	1.4	403	40	A1494731	SA14E07.Y	818	22	1.4	515	65	AA2222575
746	22	1.4	404	45	A1898615	SA268058	819	22	1.4	520	46	A1977664
747	22	1.4	405	40	A1488683	SA268058	820	22	1.4	521	47	AA283539
748	22	1.4	408	35	A1099814	SA3967.Lam	821	22	1.4	523	72	AA466480
749	22	1.4	409	39	A1437715	SA38F10.Y	822	22	1.4	527	79	AA620408
750	22	1.4	409	39	A1443076	SA47D01.Y	823	22	1.4	530	94	AQ153040
751	22	1.4	409	93	AQ074711	CIT-HSP-2	824	22	1.4	534	105	AQ662447
752	22	1.4	410	38	A1534000	SA12E84.S	825	22	1.4	534	118	AQ982815
753	22	1.4	413	39	A1442444	SA827A05.Y	826	22	1.4	535	65	AA234056
754	22	1.4	421	40	A1495640	SA1D12.Y	827	22	1.4	537	87	N8300
755	22	1.4	422	73	AA529343	UI-R-BT1	828	22	1.4	541	43	AA1724181
756	22	1.4	423	40	A1496476	SA08B11.Y	829	22	1.4	542	46	A1975307
757	22	1.4	425	40	A1495328	SA91A11.Y	830	22	1.4	543	64	AA156310
758	22	1.4	425	40	A1507820	SA87F02.Y	831	22	1.4	552	45	A1942565
759	22	1.4	427	46	A1943337	SA78H05.Y	832	22	1.4	552	80	AA715765
760	22	1.4	428	63	AA100395	SA27B12.Y	833	22	1.4	560	74	AA584177
761	22	1.4	429	63	AA052953	AA052953	834	22	1.4	560	117	AQ941909
762	22	1.4	430	118	AQ972390	RPCT-23-3	835	22	1.4	563	64	AA146679
763	22	1.4	433	93	AA064367	HS-2205.A	836	22	1.4	566	70	AA305586
764	22	1.4	434	37	A1263008	SA34E11.X	837	22	1.4	567	103	AAQ46232
765	22	1.4	436	43	A1748167	SA49E06.Y	838	22	1.4	568	122	FR0032795
766	22	1.4	436	80	AA711210	SA49E06.Y	839	22	1.4	574	79	AA685096
767	22	1.4	437	40	A1507831	SA87902.Y	840	22	1.4	575	79	AA621731
768	22	1.4	438	74	AA597697	SA49F09.Y	841	22	1.4	579	63	AA101997
769	22	1.4	439	44	A1779552	SA260431	842	22	1.4	581	42	A1641693
770	22	1.4	439	63	AA086922	SA09B06.X	843	22	1.4	581	42	A1641693
771	22	1.4	441	64	AA159411	SA98H04.X	844	22	1.4	584	118	AQ982027
772	22	1.4	444	39	A1437930	SA81E04.Y	845	22	1.4	590	105	AAQ650426
773	22	1.4	444	74	AA585534	SA731177	846	22	1.4	593	63	AA077719
774	22	1.4	445	79	AA666050	SA31E02.Y	847	22	1.4	595	118	AQ949996
775	22	1.4	445	92	AA47664	AA7S4520.St	848	22	1.4	596	74	AA586514
776	22	1.4	448	45	A1900612	SA13F08.Y	849	22	1.4	603	79	AA645402
777	22	1.4	451	35	A1104057	SA7213346	850	22	1.4	604	80	AA704425
778	22	1.4	453	29	AA678869	MA6FCW5G0	851	22	1.4	605	114	AQ766331
779	22	1.4	455	74	AA593952	SA196A03.Y	852	22	1.4	607	74	AA596167
780	22	1.4	457	63	AA102259	SA85D09.Y	853	22	1.4	619	79	AA684228
781	22	1.4	457	74	AA595957	SA7314615	854	22	1.4	620	64	AA133904
782	22	1.4	457	89	AA22782	SA46611.Y	855	22	1.4	620	70	AA283015
783	22	1.4	458	74	AA598747	SA46611.Y	856	22	1.4	620	118	AQ948293
784	22	1.4	461	113	AQ757038	HS-2244.B	857	22	1.4	627	45	A1894880
785	22	1.4	463	41	AA575193	UI-R-G0.U	858	22	1.4	628	101	AQ379440

C 859	22	1.4	632	118	A0947648	Sheared D	C 932	21	1.3	401	81	D33727	D33727 CELK03584R
C 860	22	1.4	645	73	AA497701	SRTD25CAU	C 933	21	1.3	401	120	B40339	B40339 HS-1051-A2-
C 861	22	1.4	655	73	A0662111	A0662111	C 934	21	1.3	402	126	AA409484	AA409484 EST01213
C 862	22	1.4	659	81	C90627	C90627 D1Ct	C 935	21	1.3	412	41	A1558480	A1558480 fb79g01.x
C 863	22	1.4	660	122	CNS00YD5	CNS00YD5	C 936	21	1.3	420	105	A0658202	A0658202 927P1-9H1
C 864	22	1.4	664	79	AM685572	AM685572	C 937	21	1.3	429	118	A0959390	A0959390 LER8771R
C 865	22	1.4	680	105	AO648989	AO648989	C 938	21	1.3	431	70	AM280008	AM280008 f148B08.x
C 866	22	1.4	689	79	AM692595	AM692595	C 939	21	1.3	431	90	W00775	W00775 SWAKC4579SK
C 867	22	1.4	690	95	AO250830	AO250830	C 940	21	1.3	434	34	AF164452	AF164452 B0265522
C 868	22	1.4	699	69	AM175311	AM175311	C 941	21	1.3	438	31	AA807999	AA807999 nv64a10.s
C 869	22	1.4	705	118	AO985145	AO985145	C 942	21	1.3	438	117	A0884990	A0884990 HS-5517-B
C 870	22	1.4	708	96	AO287875	AO287875	C 943	21	1.3	446	70	AM281709	AM281709 f153g12.x
C 871	22	1.4	711	48	AO057637	AO057637	C 944	21	1.3	448	41	A1601770	A1601770 fcl1e07.x
C 872	22	1.4	726	47	AO05102	AO05102	C 945	21	1.3	455	28	AA605371	AA605371 30407 Lam
C 873	22	1.4	744	69	AM223548	AM223548	C 946	21	1.3	460	89	T45370	T45370 8633 Lambda
C 874	22	1.4	764	74	AM559774	AM559774	C 947	21	1.3	462	120	B40344	B40344 HS-1051-A2-
C 875	22	1.4	770	47	AL037995	AL037995	C 948	21	1.3	465	65	AM100675	AM100675 gd58e10.y
C 876	22	1.4	772	74	AM559486	AM559486	C 949	21	1.3	466	95	AO202944	AO202944 RPT111-43
C 877	22	1.4	774	113	AO687683	AO687683	C 950	21	1.3	468	34	A1043396	A1043396 BSBMF070
C 878	22	1.4	779	71	AM331984	AM331984	C 951	21	1.3	468	46	AI976113	AI976113 EST270707
C 879	22	1.4	855	40	AL522811	AL522811	C 952	21	1.3	472	95	AO202844	AO202844 RPT111-43
C 880	22	1.4	891	122	CNS005WN	CNS005WN	C 953	21	1.3	477	45	AI883936	AI883936 fcl6h07.x
C 881	22	1.4	933	122	CNS00M15	CNS00M15	C 954	21	1.3	477	47	AL035888	AL035888 DRP2564E
C 882	22	1.4	957	122	CNS00H25	CNS00H25	C 955	21	1.3	480	105	AO664395	AO664395 HS-5480-B
C 883	22	1.4	1005	122	CNS00KOP	CNS00KOP	C 956	21	1.3	480	116	AO859663	AO859663 nbe00013I
C 884	22	1.4	1023	122	CNS0018R	CNS0018R	C 957	21	1.3	483	63	AA986631	AA986631 uc82c04.y
C 885	22	1.4	1101	122	CNS0008O	CNS0008O	C 958	21	1.3	486	33	AA053780	AA053780 L30-1674T
C 886	21	1.3	179	50	AV163887	AV163887	C 959	21	1.3	490	44	A1778368	A1778368 EST259247
C 887	21	1.3	216	48	AO060639	AO060639	C 960	21	1.3	492	121	B97033	B97033 F1N7FRB IG
C 888	21	1.3	236	37	AI253853	AI253853	C 961	21	1.3	494	21	AA103296	AA103296 mo17g10.f</

ALIGNMENTS

D33727 CELK03554R
AA03039 HS-1051-AA-
AA40948A EST01213
A1558480 5779901.x
AQ682802 527P1-9H1
AQ939930 LER8701R
AQ080908 #14808.iR
XMO0775 SWAC6519SK
PE164452 D026552Z
AA807999 nv6440.S
AQ084939 HS-3517.B
XW281709 #153912.x
YA1601370 f611e07.x
AA603371 30407 Lam
T45370 8633 lambda
B40344 HS-1051-AA-
AM100675 gds5840.y
YA1024394 RC111-43
AQ104396 BSBMMFSO
YA1976115 EST270707
AQ020844 RC111-43
YA1883936 f646007.x
AL035888 DKF25604E
AQ664336 HS-3480.B
AA983663 nbcc00131H
XAA986311 uc82c04.y
AM053780 L30-1674T
YA178366 EST252505
AM0736041 GA_EB002
AV0392121 A0039219
AL131334 Fugu_rubr
AM115657 #19810.x
YA1889554 EST266997
AQ216792 HS_2262.A
A08080249 HS_5315.A
B26300_T20N6TR.TM
YA1976124 EST270718
AW5568213 s169407.y
AM050287 EST328747.y
AA123784 mp90c02.y
XAA874482 V100927.mT
AL086753 Arabidops
AM422190 #158905.y
AQ570070 HS_5370.A
AV0534297 A0053397
YA0554293 S01033.3
A091941817 Sheared.D
AL124555 Fugu_rubr
C84683 C84633 Dict
AL002117 F.rubridipe
AM041831 EST28495E
AM091948 EST285128
AW737914 EST338756
AQ273376 g15bxd00300
AW706290 g15b4605.y
YA1258862 LP02158.3
YA1258618 LP01811.3
AW631558 EST81235E
B24345 #17197F IGF
AM001035 #191911.x
AW229021 up25c11.x
YA1582842 f0202c1.y
YA1526741 Uj42d01.y

FEATURES	RESULT	2
AM154180/c	LOCUS	AM154180/2
AM154180	DEFINITION	595 bp mRNA EST 03-NOV-1999
123020.0.x1	ACCESSION	Sugano Kawakami zebrafish DNA Danio rerio cDNA clone
2602026.3	VERSION	' similar to contains element LTR1 repetitive element ; ,
mRNA sequence.	KEYWORDS	AM154180
AM154180	SOURCE	AM154180.1 GI:6202082
EST.	ORGANISM	EST.
zebrafish.	REFERENCE	Danio rerio
Danio rerio	AUTHORS	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
1 (bases 1 to 595)	TITLE	Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S., Hillier,L., Clifton,S., Allen,M., Gibson,M., Jost,S., Kucaba,T., Martin,J., Pape,D., Steptoe,M., Underwood,K., Thelshing,B., Rutter,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
Washu Zebrafish EST Project 1999	JOURNAL	Unpublished (1999)
On Jun 5, 1998 this sequence version replaced gi:31893302.	COMMENT	Other ESTs: fl33e0.y1
Contact: S.L. Johnson		Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA		Tel: 314 286 1800
Fax: 314 286 1810		Email: est@wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNRN		Sequencing by: Washington University Genome Sequencing Center
zebrafish identity (P-value greater than 1e-99) found to:		gi:12389333 gb AA542448 AA542448 fa07a10.s1 zebrafish ICRFzfs Danio
rerio cDNA		Seq primer: T7 ET from Amersham
High quality sequence stop: 515.		Location/Qualifiers

```

FEATURES
source
location/Qualifiers
1..595
/organism="Danio rerio"
/strain="Ab"
/db_xref="taxon:7955"
/clone="2602026"
/clone_1fb="Sugano Kawakami zebrafish DRA"
/sex="mixed (one male and one female, including unfertilized eggs)"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/not_vector="PMEL18-FL3; site.1: DraIII (CACGTGCG); site.2: DraIII (CACCATGCG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor (TTGTGGCTTACTGG), digested and cloned into distinct DraIII sites of the PMEL18-FL3 vector (5' site CACTGTGCG, 3' site CACCATGCG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science) and kindly donated by Dr. Koichi Kawakami. Custom primers for sequencing: 5' end primer CTTGTGCTCTTAAAGTCGC and 3' end primer CGACCTCGACCTGCGACCA.

```

```

Query Match      2.1%; Score 33; DB 64; Length 595;
Best Local Similarity 100.0%; Pred. No. 7e-07;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 CATCATCATCATCATCATCATCATCATCATCATCAGC 39
      |||
bb 387 CATCATCATCATCATCATCATCATCATCATCATCAGC 355
      |||

```


Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CCATCATCATCATCATCATCATCATCATC 36
 DB 300 CCATCATCATCATCATCATCATCATC 270

RESULT 8
 LOCUS AQ196790/c 463 bp DNA GSS 16-SEP-1998
 DEFINITION CIT-HSP-2383019.TF CIT-HSP Homo sapiens genomic clone 2383019,
 genomic survey sequence.
 ACCESSION AQ196790
 VERSION AQ196790.1 GI:3604152
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 463)
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
 Berry,K., Granger,D., Sub,E., Wible,C., Shizuya,H., Simon,M. and
 Venter,J.C.

TITLE
 JOURNAL Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building
 COMMENT Unpublished (1998)
 Other GSS: CIT-HSP-2383019.TR
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

FEATURES
 Location/Qualifiers

1..463
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2383019"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pbe10BAC11; Site_1: HindIII; Site_2:
 HindIII"

BASE COUNT 144 a 84 c 100 g 135 t

Query Match 2.0%; Score 31; DB 95; Length 463;
 Best Local Similarity 100.0%; Pred. No. 8.7e-06;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GCCATCATCATCATCATCATCATCATCA 35
 DB 234 GCCATCATCATCATCATCATCATCA 204

RESULT 9
 LOCUS AI584760 538 bp mRNA EST 06-APR-1999
 DEFINITION ITB83609.Y1 zebrafish washu MPIMG EST Danio rerio cDNA 5', mRNA
 sequence.
 ACCESSION AI584760
 VERSION AI584760
 KEYWORDS EST.
 SOURCE zebrafish.
 ORGANISM Danio rerio

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
 1 (bases 1 to 538)
 Clark,M., Johnson,S.L., Lebrach,H., Lee,R., Li,F., Marra,M.,
 Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
 Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R.,
 Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.

TITLE
 JOURNAL Washu zebrafish EST Project 1998
 COMMENT Unpublished (1998)
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: zbrfish@watson.wustl.edu
 cDNA Library Preparation: Matthew Clark. cDNA library Arrayed by:
 Matthew Clark. DNA Sequencing by: Washington University Genome
 Sequencing Center Clone distribution: Genome Systems, St. Louis,
 Missouri (web address: www.genomesystems.com) (email contact:
 info@genomesystems.com) and Research Genetics, Huntsville, Alabama
 (web address: www.resgen.com) (email contact: info@resgen.com) and
 ResourceCenter@umr.fritzlab.tu-berlin.de, Berlin, Germany (web address:
 www.rzpd.de)

Seq primer: T3 ET from Amerham
 High quality sequence stop: 480.

FEATURES
 Location/Qualifiers

1..538
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone_lib="zebrafish Washu MPIMG EST"
 /sex="mixed"
 /tissue="mixed"
 /stage="embryos"
 /lab_host="X11-blue MRF"
 /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
 strand cDNA was primed with a Not I - oligo(dT)15 primer
 [5'-GACATGATCTAGTACGAGCGGCGCCCTTTTCTTTTCTTTT3']
 double-stranded cDNA was ligated to Sal I adaptors (BRL),
 digested with Not I and cloned into the Not I and Sal I
 sites of the pSPORT1 vector (BRL). Library was constructed
 by Matthew Clark (Lebrach lab; ICRF, London and Max Planck
 Institut fuer Molekulare Genetik, Berlin). cDNAs for EST
 analysis were selected following oligonucleotide
 hybridization fingerprinting of arrayed clones from
 zebrafish late somitogenesis (26 ss), adult liver or
 embryonic shield stage (5.6 h) libraries. Fingerprint
 data were used to computationally cluster cDNAs, and a
 single cDNA from each cluster was chosen for sequencing.
 In some cases multiple members of the same cluster were
 sequenced to assess clustering parameters or single clones
 were sequenced additional times to assess quality
 control."

BASE COUNT 157 a 88 c 118 g 174 t 1 others

Query Match 2.0%; Score 31; DB 41; Length 538;
 Best Local Similarity 100.0%; Pred. No. 8.7e-06;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CACATCATCATCATCATCATCATCATCA 37
 DB 313 CACATCATCATCATCATCATCATCA 283

RESULT 10
 LOCUS AM649105/c 545 bp mRNA EST 04-APR-2000
 DEFINITION EST327559 tomato germinating seedlings, TAMU Lycopersicon
 esculentum cDNA clone cLE17117 5', mRNA sequence.

ACCESSION AM649105
 VERSION AM649105.1 GI:7410343
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota: Viridiplantae; Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 545)
 Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nieman,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and Tanksley,S.D.
 Generation of ESTs from germinating tomato seed
 Unpublished (2000)
 On Jul 8, 1999 this sequence version replaced gi:5422601.
 Contact: David Frisch
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfrisch@CLEMSON.EDU

FEATURES
 Source
 1..545
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEI7117"
 /clone_id="tomato germinating seedlings, TAMU"
 /tissue_type="whole seedlings"
 /dev_stage="7 days post imbibition"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; 7 days post imbibition on water-agar. Mixed stage whole germinating seedlings from seed coat emergence up to two centimeters in seeds not showing obvious signs of germination were discarded."

BASE COUNT
 180 a 122 g 143 t 1 others

ORIGIN

Query Match 2.0%; Score 31; DB 79; Length 545;
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CCATCATCATCATCATCATCATCATCATCATCATC 36
 |||||||||||||||||||||||||||||||||||
 DB 520 CCATCATCATCATCATCATCATCATCATCATC 490

RESULT 11
 AM624043/c 600 bp mRNA EST 28-MAR-2000
 LOCUS EST3321988 tomato flower buds 3-8 mm, Cornell University
 DEFINITION Lycopersicon esculentum cDNA clone CTOB13N16 5', mRNA sequence.
 ACCESSION AM624043
 VERSION AM624043.1 GI:7337070
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota: Viridiplantae; Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 600)
 van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nieman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
 Generation of ESTs from tomato flower tissue, 3-8 mm buds
 Unpublished (1999)
 On Dec 5, 1997 this sequence version replaced gi:2662833.
 Contact: David Frisch

TITLE
 JOURNAL
 COMMENT

ACCESSION AM649317
 VERSION AM649317.1 GI:7410555
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota: Viridiplantae; Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 628)
 Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nieman,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and Tanksley,S.D.
 Generation of ESTs from germinating tomato seed
 Unpublished (2000)
 On Jan 6, 2000 this sequence version replaced gi:6675154.
 Contact: David Frisch
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfrisch@CLEMSON.EDU

FEATURES
 Source
 1..628
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEI732"
 /clone_id="tomato germinating seedlings, TAMU"
 /tissue_type="whole seedlings"
 /dev_stage="7 days post imbibition"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Tanksley; flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT
 203 a 100 c 132 g 164 t 1 others

ORIGIN

Query Match 2.0%; Score 31; DB 79; Length 600;
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CCATCATCATCATCATCATCATCATCATCATCATC 36
 |||||||||||||||||||||||||||||||||||
 DB 378 CCATCATCATCATCATCATCATCATCATCATC 348

RESULT 12
 AM649317/c 628 bp mRNA EST 04-APR-2000
 LOCUS EST327771 tomato germinating seedlings, TAMU Lycopersicon
 DEFINITION esculentum cDNA clone CLEI732 5', mRNA sequence.
 ACCESSION AM649317
 VERSION AM649317.1 GI:7410555
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota: Viridiplantae; Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 628)
 Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nieman,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and Tanksley,S.D.
 Generation of ESTs from germinating tomato seed
 Unpublished (2000)
 On Jan 6, 2000 this sequence version replaced gi:6675154.
 Contact: David Frisch
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfrisch@CLEMSON.EDU

TITLE
 JOURNAL
 COMMENT

Whole, 7 days post imbibition on water-agar. Mixed stage whole germinating seedlings from seed coat emergence up to two centimeters in seeds not showing obvious signs of germination were discarded."

BASE COUNT 209 a 115 c 142 g 161 t 1 others

Query Match 2.0%; Score 31; DB 79; Length 628;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CCATCATCATCATCATCATCATCATCATCATC 36
DB 518 CCATCATCATCATCATCATCATCATCATCATC 488

RESULT 13

LOCUS C24646 650 bp mRNA EST 28-APR-1999
DEFINITION C24646 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium discoideum cDNA clone SL-X046, mRNA sequence.

ACCESSION C24646.1 GI:2243067
VERSION C24646.1
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum.
REFERENCE Dictyostelium; Dictyostellida; Dictyostelium.
AUTHORS Morio, T., Urushihara, H., Salto, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Mita, B. N., Pl. M., Salto, T., Takemoto, K., Yasukawa, H., Williams, D., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
The Dictyostelium developmental CDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development

JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99136227
COMMENT On Jul 9, 1999 this sequence version replaced g1:5435108.
Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Tennodai, Tsukuba, Ibaraki 305, Japan
Email: d40zhuesakura.cc.tsukuba.ac.jp
PROJECT - 'Dictyostelium discoideum CDNA project in Japan'
POLY-A-No.

FEATURES Location/Qualifiers

1..650
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SL-X046"
/clone_1lb="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"

BASE COUNT 231 a 92 c 124 g 203 t

Query Match 2.0%; Score 31; DB 80; Length 650;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CATCATCATCATCATCATCATCATCATCATC 37
DB 40 CATCATCATCATCATCATCATCATCATCATC 70

RESULT 14

LOCUS AQ349263 113 bp DNA GSS 07-MAY-1999
DEFINITION RPCI11-139F19, TV RPCI-11 Homo sapiens genomic clone RPCI-11-139F19, genomic survey sequence.
ACCESSION AQ349263

VERSION AQ349263.1 GI:4174159
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 113)
Zao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@tigr.org, med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: 17
Class: BAC ends.

FEATURES Location/Qualifiers

1..113
/organism="Homo sapiens"
/db_xref="GDB:755130"
/db_xref="taxon:9606"
/clone="RPCI-11-139F19"
/clone_1lb="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site: 1: EcoRI; Site: 2: EcoRI; RPCI11 Human Male BAC Library"

BASE COUNT 33 a 23 c 17 g 40 t
Query Match 1.9%; Score 30; DB 101; Length 113;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ATCATCATCATCATCATCATCATCATCATCATC 37
DB 9 ATCATCATCATCATCATCATCATCATCATCATC 38

RESULT 15
LOCUS HS021500 198 bp DNA GSS 10-MAR-1997
DEFINITION Human chromosome 17q21 clone ET10, genomic survey sequence.
ACCESSION U21500
VERSION U21500.1 GI:1872174
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 198)
McFarlane, R., de Jong, P., Quirk, V.P., Heinrich, H. and Solomon, E.
Physical mapping, cloning, and identification of genes within a 500-kb region containing BRCA1
Proc. Natl. Acad. Sci. U.S.A. 92 (10), 4362-4366 (1995)
95273363
2 (bases 1 to 198)
Brown, M.A.
Direct Submission
Submitted (22-FEB-1995) Melissa A. Brown, Somatic Cell Genetics,

REFERENCE Brown, M.A., Jones, K.A., Nicolai, H., Bonjardim, M., Black, D., McFarlane, R., de Jong, P., Quirk, V.P., Heinrich, H. and Solomon, E.
Physical mapping, cloning, and identification of genes within a 500-kb region containing BRCA1
Proc. Natl. Acad. Sci. U.S.A. 92 (10), 4362-4366 (1995)
JOURNAL 95273363
MEDLINE 95273363
REFERENCE 2 (bases 1 to 198)
Brown, M.A.
Direct Submission
Submitted (22-FEB-1995) Melissa A. Brown, Somatic Cell Genetics,

Imperial Cancer Research Fund, Lincoln's Inn Fields, London, WC2A
3PX, UK
COMMENT On Mar 10, 1997 this sequence version replaced g1:852476.
FEATURES Location/Qualifiers
Source 1..198
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ET10"
/chromosome="17"
/map="17q21"
/note="exon-trapped clone isolated using the procedure of
Buckler et al. PNAS (1991) 88:4005-4009"

BASE COUNT 54 a 28 c 58 g 58 t
ORIGIN

Query Match 1.9% Score 30; DB 123; Length 198;
Best Local Similarity 100.0%; Pred.No. 2.9e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 CCATCATCATCATCATCATCATCA 35
|||||
DB 129 CCATCATCATCATCATCATCATCA 100

Search completed: August 22, 2000, 16:04:40
Job time: 3243 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:02:20 ; Search time 29.04 Seconds
(without alignments)
424.947 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGNHHHHHHSSGHIDD.....RFIOKYDSVMQRLIDPTSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 188963 seqs, 23686106 residues

Word size: 0

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database: A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	46.6	329	1 R79961	Partial LcrV (V an
2	243	46.6	329	1 W01040	Y. pestis V antige
3	243	46.6	329	1 W01041	Y. pestis V antige
4	207	39.7	480	1 W01045	Y. pestis Fl/V ant
5	170	32.6	170	1 R76528	Yersinia pestis ca
6	170	32.6	170	1 W01043	Y. pestis Fl antlg
7	170	32.6	170	1 W59782	Amino acid sequenc
8	170	32.6	501	1 W01044	Y. pestis Fl/V ant
9	170	32.6	329	1 R79962	Partial LcrV (V an
10	167	33.1	329	1 W59787	Nucleotide sequenc
11	149	28.6	149	1 W59787	Amino acid sequenc
12	149	28.6	150	1 W59787	Amino acid sequenc
13	134	22.7	192	1 W59785	Yersinia pestis ca
14	128	24.6	151	1 R76526	Yersinia pestis ca
15	128	24.6	151	1 R76527	Y. pestis Fl antlg
16	128	24.6	151	1 W01042	Amino acid sequenc
17	113	21.7	171	1 W59786	Y. pestis Fl antlg
18	24	4.6	24	1 W19779	Amino acid sequenc
19	17	3.3	33	1 W36986	Polyhistidine-ente
20	17	3.3	54	1 W36990	HFR1 peptide. Meth
21	17	3.3	323	1 W71185	FP505 protein cont
22	17	3.3	423	1 W68098	Chlamydomonas rein
23	17	3.3	448	1 W68395	Clostridium botuli
24	17	3.3	451	1 W68395	Clostridium botuli
25	17	3.3	451	1 W68398	Clostridium botuli
26	17	3.3	452	1 W68396	Clostridium botuli
27	17	3.3	462	1 R95009	Type A neurotoxin
28	17	3.3	462	1 W68390	Clostridium botuli
29	17	3.3	462	1 W68397	Clostridium botuli
30	17	3.3	472	1 W68394	Clostridium botuli
31	17	3.3	472	1 W68393	Clostridium botuli
32	17	3.3	473	1 W68400	Clostridium botuli
33	16	3.1	21	1 W23647	Recombinant squitr

34	16	3.1	21	1 W23650	Recombinant squitr
35	16	3.1	21	1 W23653	Recombinant squitr
36	16	3.1	254	1 W37699	Protein sequence o
37	12	2.3	17	1 W39399	T2 peptide, synthe
38	11	2.1	302	1 R82097	Human Gax protein.
39	11	2.1	303	1 R82096	Rat Gax protein. R
40	11	2.1	414	1 R32020	Sequence of a euka
41	11	2.1	414	1 W65406	Human transcriptio
42	10	1.9	96	1 W99830	HIV Var protein se
43	10	1.9	96	1 W99831	HIV E21,24P protei
44	10	1.9	96	1 W99832	HIV A30S protein s
45	10	1.9	96	1 W99833	HIV A30L protein s
46	10	1.9	96	1 W99834	HIV A59P protein s
47	10	1.9	96	1 W99835	HIV L68S protein s
48	10	1.9	96	1 W99836	HIV H71C protein s
49	10	1.9	96	1 W99837	HIV G75S protein s
50	10	1.9	96	1 W99838	HIV C76S protein s
51	10	1.9	335	1 W94462	Human wild-type Ho
52	10	1.9	351	1 R24393	Sequence of Histid
53	10	1.9	449	1 W56703	Human GATA-6 trans
54	10	1.9	480	1 R44551	Brain factor-1. Is
55	10	1.9	626	1 W16398	Human neuron-deriv
56	10	1.9	628	1 R32057	Apoptotic cerebral
57	10	1.9	672	1 R31216	Penicillin binding
58	10	1.9	763	1 W41734	Human TRAF-2 kinas
59	10	1.9	786	1 R47066	Sequence of Crypto
60	10	1.9	2262	1 W56737	Calcium ion channe
61	10	1.9	2266	1 R11008	Human neuronal cal
62	10	1.9	2510	1 R71007	Human neuronal cal
63	9	1.7	18	1 W73374	Epitope tagged TBP
64	9	1.7	19	1 W00339	Rec2 protein leade
65	9	1.7	20	1 W08594	Leader sequence fo
66	9	1.7	21	1 R87745	Hexa-His/thrombin
67	9	1.7	21	1 W08590	Leader sequence fo
68	9	1.7	21	1 W97132	Fusion peptide lea
69	9	1.7	22	1 W35362	Synthetic 65K-gut
70	9	1.7	23	1 W65551	peptide expressed
71	9	1.7	25	1 R92906	His-Tag leader seq
72	9	1.7	43	1 R82733	PET-15b expression
73	9	1.7	110	1 W78908	Rat CARK1 fusion p
74	9	1.7	137	1 R55693	hnRNP U protein 69
75	9	1.7	144	1 W47195	Heppes simplex vir
76	9	1.7	166	1 W00525	Human obesity prot
77	9	1.7	166	1 W00529	Human obesity prot
78	9	1.7	166	1 W00533	Human obesity prot
79	9	1.7	166	1 W00524	Human obesity prot
80	9	1.7	167	1 R92730	Human obesity prot
81	9	1.7	167	1 W00521	Human obesity prot
82	9	1.7	167	1 R92726	Human obesity prot
83	9	1.7	167	1 W00527	Human obesity prot
84	9	1.7	180	1 W47081	Salmonella Sef1 p
85	9	1.7	359	1 W17791	Maize nuclear-loca
86	9	1.7	359	1 W17794	KN1 alanine scann
87	9	1.7	359	1 W17795	KN1 alanine scann
88	9	1.7	359	1 W17796	KN1 alanine scann
89	9	1.7	359	1 W17797	KN1 alanine scann
90	9	1.7	359	1 W17798	KN1 alanine scann
91	9	1.7	359	1 W17799	KN1 alanine scann
92	9	1.7	359	1 W17800	KN1 alanine scann
93	9	1.7	371	1 W73369	Epitope tagged TBP
94	9	1.7	398	1 W71071	Multiple sclerosis
95	9	1.7	398	1 W99557	Protein encoded by
96	9	1.7	402	1 R24392	Sequence of the Hi
97	9	1.7	408	1 W84315	TrxA-rabbit tissue
98	9	1.7	421	1 W68541	Amino acid sequenc
99	9	1.7	423	1 W96263	Bn-36 polypeptide
100	9	1.7	423	1 W96262	Topoisomerase I ex
101	9	1.7	711	1 W57321	Protein allergen I
102	8	1.5	10	1 W72331	N-terminal peptide
103	8	1.5	23	1 W15248	Nickel binding pro
104	8	1.5	60	1 R90286	H. pylori cytoplas
105	8	1.5	60	1 W20250	H. pylori cytoplas
106	8	1.5	60	1 W24617	H. pylori cytoplas

107	1.5	64	1	W27485	Sequence used in d
108	1.5	78	1	W20756	H. pylori cytoplas
109	1.5	96	1	R36548	Recombinant Y2X. R
110	1.5	221	1	W23460	Recombinant MAM su
111	1.5	335	1	W94465	Human HoxA1 varian
112	1.5	403	1	W72943	Mycobacterium tube
113	1.5	404	1	W72942	Mycobacterium tube
114	1.5	494	1	W23066	Rabbit beta-2 inte
115	1.5	494	1	W65108	Rabbit beta-2 inte
116	1.5	494	1	W72840	Rabbit alpha-d clo
117	1.5	494	1	W73350	Rabbit alpha-d clo
118	1.5	531	1	W47002	Glutathione-S-tran
119	1.5	536	1	R88597	Fork insertion mut
120	1.5	600	1	R88598	Fork insertion mut
121	1.5	802	1	W64379	Mycobacterium anti
122	1.5	802	1	W81746	M. tuberculosis fu
123	1.3	9	1	W06142	Variant adenovirus
124	1.3	13	1	R72792	Hexa-histidine lea
125	1.3	17	1	R68877	pGEX-7Hpro proteas
126	1.3	73	1	Y01320	Human secreted pro
127	1.3	73	1	Y01331	Modified K11 RNA p
128	1.3	73	1	Y01295	K11 RNA polymerase
129	1.3	73	1	Y01362	Modified K11 RNA p
130	1.3	84	1	W00945	CMV500-4heptadREB
131	1.3	91	1	W46607	Human brain protei
132	1.3	96	1	W38487	Streptococcus pneu
133	1.3	97	1	W00949	CMV500-4heptadfos
134	1.3	105	1	W00947	CMV500-Fosb2IP(MO)
135	1.3	111	1	W59030	Carbomaterioclin B
136	1.3	111	1	W94888	Carbomaterioclin B
137	1.3	113	1	W00950	CMV400-4undzip leu
138	1.3	124	1	W78909	Rat CAR2 fusion p
139	1.3	129	1	W54310	Plasmid DNA encodi
140	1.3	139	1	W54285	Human modified cyt
141	1.3	135	1	W59053	Hs rFGF10 protein
142	1.3	136	1	W71694	R21-2Hpro-28 fusio
143	1.3	210	1	R85066	Wilms' tumour prot
144	1.3	210	1	W22883	Wilms' tumour prot
145	1.3	220	1	Y02631	Prot.DI/3-E7-mut(C
146	1.3	220	1	Y02634	Prothrlh126-E7-His
147	1.3	227	1	Y02638	Prot.DI/3-E7-His/H
148	1.3	227	1	Y02640	Prot.DI/3-E7-mut(C
149	1.3	239	1	Y02636	CLYRA-E7-His prote
150	1.3	249	1	W06148	Synthetic adenovir
151	1.3	273	1	Y02632	Prot.DI/3-E6-His/H
152	1.3	278	1	Y02641	Prot.DI/3-E6-His/H
153	1.3	282	1	Y02635	CLYRA-E6-His prote
154	1.3	302	1	Y05097	HIV Tat protein. p
155	1.3	349	1	W54311	Pleckstrin homology
156	1.3	371	1	Y02633	Prot.DI/3-E6-E7-Hi
157	1.3	383	1	Y02642	Prot.DI/3-E6-E7-Hi
158	1.3	390	1	Y02637	CLYRA-E6E7-His pro
159	1.3	391	1	W23534	CDK inhibitory fus
160	1.3	381	1	W95094	Human p27-p16 fusi
161	1.3	400	1	R75203	Tyrosine phosphata
162	1.3	433	1	W17783	FIV integrase-Lexa
163	1.3	433	1	W17784	FIV integrase-Lexa
164	1.3	515	1	W55073	Streptococcus pneu
165	1.3	585	1	R88595	Fork insertion mut
166	1.3	589	1	R88596	Fork insertion mut
167	1.3	685	1	W88432	Disease associated
168	1.3	685	1	Y00915	Human serum induci
169	1.3	713	1	W21960	Recombinant furin
170	1.3	806	1	R65495	Marek's disease vi
171	1.3	1262	1	W13505	B. bronchiseptica
172	1.3	1263	1	W13503	B. pertussis adeny
173	1.3	1445	1	R12108	N-terminal deleted
174	1.3	1522	1	P93357	Sequence of the ca
175	1.3	1644	1	W13504	B. bronchiseptica
176	1.3	1645	1	W13502	B. pertussis adeny
177	1.3	1705	1	P94365	Sequence of part o
178	1.3	1706	1	R08031	Adenyl cyclase fro
179	1.3	1891	1	W52844	Amycolatopsis medi
180	1.3	3413	1	W52849	A. mediterranei ri
181	1.3	3729	1	W22603	Tyactone synthase
182	1.2	6	1	W19161	Isoelectric point
183	1.2	6	1	W18225	Purification tag o
184	1.2	6	1	W44011	Poly-histidine pep
185	1.2	6	1	W69961	Poly-His tag for c
186	1.2	6	1	W68297	Poly-His tag for c
187	1.2	6	1	W63024	Hexa-histidine tag
188	1.2	6	1	W84203	Peptide comprising
189	1.2	7	1	P80401	Affinity peptide c
190	1.2	7	1	Y03173	Human cardiac trop
191	1.2	8	1	R77412	Extension peptide
192	1.2	8	1	R95439	Polyhistidine tail
193	1.2	8	1	W20036	C-terminal metal b
194	1.2	8	1	W61479	A. fumigatus aller
195	1.2	9	1	R97378	Penta-histidine pe
196	1.2	9	1	W60080	Homo sapiens RH ep
197	1.2	9	1	W90199	B7.2-CHis tag fusi
198	1.2	10	1	R27517	Effector cell prot
199	1.2	10	1	R39733	First type QE-C-t
200	1.2	10	1	R73684	Metal-affinity dec
201	1.2	10	1	R74672	Polyhistidine moti
202	1.2	10	1	R74669	I-domain N-termina
203	1.2	10	1	W31496	Human Dnae II pro
204	1.2	10	1	W36089	Human monocycle che
205	1.2	10	1	W76939	Fusion immunoglobu
206	1.2	10	1	W82996	Human fibronectin
207	1.2	10	1	W89262	N-terminal hexahis
208	1.2	10	1	W94259	His-tag sequence.
209	1.2	10	1	Y04360	Salmonella flagell
210	1.2	11	1	R77410	Extension peptide
211	1.2	11	1	W53171	C-terminal peptide
212	1.2	11	1	W61480	A. fumigatus aller
213	1.2	11	1	W70581	Amino acid sequenc
214	1.2	11	1	W80426	Linker used in the
215	1.2	12	1	R07064	Transcript of plas
216	1.2	12	1	R45783	Farnesyltransferas
217	1.2	12	1	W61549	Endoproteinase Xa
218	1.2	13	1	W82993	Human fibronectin
219	1.2	13	1	W61548	Endoproteinase Xa
220	1.2	14	1	R60018	Can fi peptide wit
221	1.2	14	1	W39836	Light chain CDRI o
222	1.2	14	1	W62982	Peptide inhibitor
223	1.2	15	1	R60513	Hexahistidine-cont
224	1.2	15	1	W72895	Mycobacterium tube
225	1.2	16	1	R69813	Cleavage site reco
226	1.2	16	1	W07945	gp120 peptide 17.
227	1.2	16	1	W30929	Birch pollen aller
228	1.2	17	1	P92081	Peptide coded for
229	1.2	17	1	R37708	Delta14 Ser17 hCNT
230	1.2	17	1	R97389	CC49 VH-spacer-PLA
231	1.2	17	1	W95000	N-terminal histidi
232	1.2	18	1	R60253	Vector-encoded NF-
233	1.2	18	1	R68814	Site recognised En
234	1.2	18	1	R78152	Synthetic leader t
235	1.2	18	1	R95870	CR2 receptor ligand
236	1.2	18	1	W71424	Peptide having an
237	1.2	19	1	W37707	C-terminal additio
238	1.2	19	1	W98067	Streptococcus pyog
239	1.2	20	1	R73014	N-terminal peptide
240	1.2	20	1	W00869	T-cell receptor al
241	1.2	20	1	W01445	Bioelastic polymer
242	1.2	20	1	W76938	Fusion immunoglobu
243	1.2	21	1	R07063	Transcript of plas
244	1.2	21	1	P71703	Internal fragment
245	1.2	21	1	R60142	pGEX-6Hpro proteas
246	1.2	21	1	W48442	Myc-his peptide ta
247	1.2	21	1	W54281	Myc-histidine pep
248	1.2	21	1	W54435	Human pS112 expres
249	1.2	21	1	W75875	Myc-his tag peptid
250	1.2	21	1	W79683	Human CS198 protei
251	1.2	21	1	W86030	Myc-His peptide ta
252	1.2	21	1	W95563	Myc-his tag used i

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:02:20 ; Search time 29.04 Seconds

(without alignments)
424.947 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGNHHHHHHSSGHIDDD.....RFIQKDYDVMQRLDPTSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 188963 seqs, 23686106 residues

Word size : 0

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	46.6	329	1 R79961	Partial LcrV (V an
2	243	46.6	329	1 W01040	Y. pestis V antige
3	243	46.6	329	1 W01041	Y. pestis V antige
4	207	39.7	480	1 W01045	Y. pestis Fl/V ant
5	170	32.6	170	1 R76528	Yersinia pestis ca
6	170	32.6	170	1 W01043	Y. pestis Fl antig
7	170	32.6	170	1 W59782	Amino acid sequenc
8	170	32.6	170	1 W59783	Amino acid sequenc
9	170	32.6	501	1 W01044	Y. pestis Fl/V ant
10	167	32.1	329	1 R79962	Partial LcrV (V an
11	149	28.6	149	1 W59788	Nucleotide sequenc
12	149	28.6	150	1 W59787	Amino acid sequenc
13	134	25.7	192	1 W59785	Yersinia pestis ca
14	128	24.6	151	1 R76526	Yersinia pestis ca
15	128	24.6	151	1 R76527	Yersinia pestis ca
16	128	24.6	151	1 W01042	Y. pestis Fl antig
17	113	21.7	171	1 W59786	Amino acid sequenc
18	24	4.6	24	1 W18779	Polyhistidine-ente
19	17	3.3	53	1 W36986	HFR1 peptide. Meth
20	17	3.3	54	1 W36990	FP505 protein cont
21	17	3.3	323	1 W71185	Chlamydomonas rein
22	17	3.3	423	1 W68098	Chlamydomonas rein
23	17	3.3	448	1 W68399	Clostridium botuli
24	17	3.3	451	1 W68399	Clostridium botuli
25	17	3.3	451	1 W68399	Clostridium botuli
26	17	3.3	452	1 W68396	Clostridium botuli
27	17	3.3	462	1 R95009	Clostridium botuli
28	17	3.3	462	1 W68390	Type A neurotoxin
29	17	3.3	462	1 W68397	Clostridium botuli
30	17	3.3	472	1 W68394	Clostridium botuli
31	17	3.3	472	1 W68393	Clostridium botuli
32	17	3.3	473	1 W68400	Clostridium botuli
33	16	3.1	21	1 W23647	Recombinant squitr

34	16	3.1	21	1 W23650	Recombinant squitr
35	16	3.1	21	1 W23653	Recombinant squitr
36	16	3.1	254	1 W37699	Protein sequence o
37	12	2.3	17	1 W03939	T2 peptide, synthe
38	11	2.1	302	1 R82097	Human Gax protein.
39	11	2.1	303	1 R82096	Rat Gax protein. R
40	11	2.1	303	1 R82096	Sequence of a euka
41	11	2.1	414	1 W65406	Human transcriptio
42	10	1.9	96	1 W99830	HIV Vpr protein se
43	10	1.9	96	1 W99831	HIV E2L24P protei
44	10	1.9	96	1 W99832	HIV A30S protein s
45	10	1.9	96	1 W99833	HIV A30L protein s
46	10	1.9	96	1 W99834	HIV A59P protein s
47	10	1.9	96	1 W99835	HIV L68S protein s
48	10	1.9	96	1 W99836	HIV H71C protein s
49	10	1.9	96	1 W99837	HIV G76S protein s
50	10	1.9	96	1 W99838	HIV C76S protein s
51	10	1.9	335	1 W94462	Human wild-type Ho
52	10	1.9	351	1 R24393	Sequence of Histid
53	10	1.9	449	1 W56703	Human GATRA-6 trans
54	10	1.9	480	1 R44551	Brain factor-1. Is
55	10	1.9	626	1 W16398	Human neuron-deriv
56	10	1.9	628	1 R32057	Apoptotic cerebral
57	10	1.9	672	1 R31216	Penicillin binding
58	10	1.9	763	1 W41734	Human TRAF-2 kinas
59	10	1.9	786	1 R47066	Sequence of Crypto
60	10	1.9	2262	1 W56737	Calcium ion channe
61	10	1.9	2266	1 R71008	Human neuronal cal
62	10	1.9	2510	1 R71007	Human neuronal cal
63	9	1.7	18	1 W73374	Epitope tagged TBP
64	9	1.7	19	1 W00339	Rec2 protein leade
65	9	1.7	20	1 W08594	Leader sequence fo
66	9	1.7	21	1 R87745	Hexa-His/thrombin
67	9	1.7	21	1 W08590	Leader sequence fo
68	9	1.7	21	1 W97132	Fusion peptide lea
69	9	1.7	22	1 W35362	Synthetic 65K-glt
70	9	1.7	23	1 W65551	Peptide expressed
71	9	1.7	25	1 R92906	His-Tag leader seq
72	9	1.7	43	1 R92733	PER-1b expression
73	9	1.7	110	1 W78908	Rat CART1 fusion p
74	9	1.7	137	1 R56993	Hepes simplex vir
75	9	1.7	144	1 W47195	Human obesity prot
76	9	1.7	166	1 W00525	Human obesity prot
77	9	1.7	166	1 W00529	Human obesity prot
78	9	1.7	166	1 W00533	Human obesity prot
79	9	1.7	166	1 W00524	Human obesity prot
80	9	1.7	167	1 R92730	Human obesity prot
81	9	1.7	167	1 W00521	Human obesity prot
82	9	1.7	167	1 R92726	Human obesity prot
83	9	1.7	167	1 W00517	Human obesity prot
84	9	1.7	180	1 W47081	Human obesity prot
85	9	1.7	359	1 W17791	Salmonella Sef1a p
86	9	1.7	359	1 W17794	Maize nuclear-10ca
87	9	1.7	359	1 W17795	KNI alanine scann
88	9	1.7	359	1 W17796	KNI alanine scann
89	9	1.7	359	1 W17797	KNI alanine scann
90	9	1.7	359	1 W17798	KNI alanine scann
91	9	1.7	359	1 W17799	KNI alanine scann
92	9	1.7	371	1 W19600	KNI alanine scann
93	9	1.7	371	1 W73369	KNI alanine scann
94	9	1.7	398	1 W71071	Epitope tagged TBP
95	9	1.7	398	1 W99557	Multiple sclerosis
96	9	1.7	402	1 R24392	Protein encoded by
97	9	1.7	420	1 W68541	Sequence of the H1
98	9	1.7	420	1 W68541	Trxa-rabbit tissue
99	9	1.7	421	1 W96263	Amino acid sequenc
100	9	1.7	423	1 W96262	Bm-3a polypeptide
101	9	1.7	711	1 W57321	Topoisomerase I ex
102	8	1.5	10	1 W72331	Protein allergen 1
103	8	1.5	23	1 W15248	N-terminal peptide
104	8	1.5	60	1 R90286	Nickel binding pro
105	8	1.5	60	1 W20250	H. pylori cytoplas
106	8	1.5	60	1 W24617	H. pylori cytoplas

107	8	1.5	64	1	W27485	Sequence used in d	180	1.3	3413	1	W52849	A. mediterranei r1
108	8	1.5	78	1	W20756	R. p10r1 cytoplas	181	1.3	3729	1	W22603	tylactone synthase
109	8	1.5	96	1	R36548	Recombinant YZX. R	182	1.2	6	1	W19161	isolectin point
110	8	1.5	221	1	W23460	Recombinant MAM su	183	1.2	6	1	W18225	purification tag
111	8	1.5	335	1	W94465	Human Hoxa1 varian	184	1.2	6	1	W44011	poly-histidine pep
112	8	1.5	403	1	W72943	Mycobacterium tube	185	1.2	6	1	W69961	poly-His tag for c
113	8	1.5	404	1	W72942	Mycobacterium tube	186	1.2	6	1	W68297	poly-His tag for c
114	8	1.5	494	1	W23066	Rabbit beta 2 inte	187	1.2	6	1	W63024	Hexa-histidine tag
115	8	1.5	494	1	W65108	Rabbit beta-integr	188	1.2	6	1	W84203	peptide comprising
116	8	1.5	494	1	W72840	Rabbit alpha-d clo	189	1.2	7	1	P80401	Affinity peptide c
117	8	1.5	494	1	W73350	Rabbit alpha-d clo	190	1.2	7	1	Y03173	Human cardiac trop
118	8	1.5	531	1	W47002	Glutathione-S-tran	191	1.2	8	1	R77412	Extension peptide
119	8	1.5	596	1	R88597	Forki insertion mut	192	1.2	8	1	R95439	Polyhistidine tail
120	8	1.5	600	1	R88598	Forki insertion mut	193	1.2	8	1	W20036	C-terminal metal b
121	8	1.5	802	1	W64379	M. tuberculosis fu	194	1.2	8	1	W61479	A. fumigatus aller
122	8	1.5	802	1	W81746	M. tuberculosis fu	195	1.2	9	1	R97378	Penta-histidine pe
123	7	1.3	9	1	W06142	Variant adenovirus	196	1.2	9	1	W60080	Human sialins RH ep
124	7	1.3	13	1	R72792	Hexa-histidine lea	197	1.2	9	1	W90199	B7.2-His tag fusi
125	7	1.3	17	1	R68977	pEX-7HPro proteas	198	1.2	10	1	R27517	Effector cell prot
126	7	1.3	64	1	Y00330	Human secreted pro	199	1.2	10	1	R39733	First type QE-C-t
127	7	1.3	73	1	Y01331	Modified K11 RNA p	200	1.2	10	1	R73684	Metal-affinity dec
128	7	1.3	73	1	Y01295	K11 RNA polymerase	201	1.2	10	1	R74672	Polyhistidine moti
129	7	1.3	73	1	Y01362	Modified k11 RNA p	202	1.2	10	1	R77469	I-domain N-termina
130	7	1.3	84	1	W00945	CMV500-4heptadcreb	203	1.2	10	1	W31496	Human DNase II pro
131	7	1.3	91	1	W46607	Human brain protei	204	1.2	10	1	W56089	Fusion monocycle che
132	7	1.3	96	1	W38487	Streptococcus pneu	205	1.2	10	1	W76939	Human fibronectin
133	7	1.3	97	1	W00949	CMV500-4heptadfos	206	1.2	10	1	W82962	N-terminal hexahis
134	7	1.3	105	1	W00947	CMV500-FosB2P(MO)	207	1.2	10	1	W82962	His-tag sequence.
135	7	1.3	111	1	W59030	Carnobacteriocin B	208	1.2	10	1	W94259	Salmonella flagell
136	7	1.3	111	1	W94888	Carnobacteriocin B	209	1.2	10	1	Y04360	Extension peptide
137	7	1.3	113	1	W00950	CMV400-dunb2ip leu	210	1.2	11	1	R77410	C-terminal peptide
138	7	1.3	124	1	W78909	Rat CAR2 fusion p	211	1.2	11	1	W53171	A. fumigatus aller
139	7	1.3	139	1	W54310	Plasmid DNA encodi	212	1.2	11	1	W61480	Transcript of plas
140	7	1.3	139	1	W54285	Human modified cyt	213	1.2	11	1	W70581	Endoproteinase Xa
141	7	1.3	135	1	W59053	His REGF10 protein	214	1.2	11	1	W80426	Endoproteinase Xa
142	7	1.3	136	1	W71684	R21-2Hpro-28 fusio	215	1.2	12	1	R07064	Human fibronectin
143	7	1.3	210	1	R85066	Wilms' tumour anti	216	1.2	12	1	R45783	Can fit peptide vlt
144	7	1.3	210	1	W22883	Wilms' tumour prot	217	1.2	12	1	W61549	Light chain CDRI O
145	7	1.3	220	1	Y02634	Prot.DI/3-E7-mut(C	218	1.2	12	1	W82993	Peptide inhibitor
146	7	1.3	220	1	Y02631	Prot.DI/3-E7-mut(C	219	1.2	13	1	W61548	Hexahistidine-cont
147	7	1.3	227	1	Y02638	Prot.DI/3-E7-His/H	220	1.2	14	1	R60018	Mycobacterium tube
148	7	1.3	227	1	Y02640	Prot.DI/3-E7-mut(C	221	1.2	14	1	W39836	Cleavage site reco
149	7	1.3	239	1	Y02636	CLYTA-E7-His prote	222	1.2	14	1	W62982	gpi20 peptide 17.
150	7	1.3	239	1	W06148	Synthetic adenovir	223	1.2	15	1	W72895	Birch pollen aller
151	7	1.3	243	1	Y02632	Prot.DI/3-E6-His/H	224	1.2	15	1	R60513	Peptide coded for
152	7	1.3	278	1	Y02641	Prot.DI/3-E6-His/H	225	1.2	16	1	R69813	Delta14 Ser17 hCNT
153	7	1.3	292	1	Y02635	CLYTA-E6-His prote	226	1.2	16	1	W07945	CC49 VH-spacer-PUA
154	7	1.3	302	1	Y05097	HIV Tat protein. p	227	1.2	16	1	W30929	N-terminal histidi
155	7	1.3	349	1	W54311	Plecksstrin homolo	228	1.2	17	1	P92081	Vector-encoded NF-
156	7	1.3	371	1	Y02633	Prot.DI/3-E6-E7-H1	229	1.2	17	1	R37708	Site recognised En
157	7	1.3	383	1	Y02642	Prot.DI/3-E6-E7-H1	230	1.2	17	1	R97389	Synthetic leader t
158	7	1.3	390	1	Y02637	CLYTA-E6E7-His pro	231	1.2	17	1	W95000	CR2 receptor ligand
159	7	1.3	391	1	W23534	CDK inhibitory fus	232	1.2	18	1	R60253	Peptide having an
160	7	1.3	391	1	W95084	Human p27-p16 fusi	233	1.2	18	1	R62914	C-terminal additio
161	7	1.3	400	1	R75203	Tyrosine phosphat	234	1.2	18	1	R78152	Streptococcus pyog
162	7	1.3	453	1	W17783	FIV integrase-Lexa	235	1.2	18	1	R95870	N-terminal peptide
163	7	1.3	499	1	W17784	RTV integrase-Lexa	236	1.2	18	1	W71424	T-cell receptor al
164	7	1.3	515	1	W55073	Streptococcus pneu	237	1.2	19	1	W37707	Bioelastic polymer
165	7	1.3	585	1	R88595	Forki insertion mut	238	1.2	19	1	W98067	Fusion immunoglob
166	7	1.3	685	1	R88452	Disease associated	239	1.2	20	1	R73014	Transcript of plas
167	7	1.3	685	1	Y00915	Human serum induc	240	1.2	20	1	W00869	Internal fragment
168	7	1.3	685	1	W21960	Recombinant furin	241	1.2	20	1	W01445	pEX-6Hpro proteas
169	7	1.3	806	1	R65495	Marek's disease vl	242	1.2	20	1	W76938	Myc-his peptide ta
170	7	1.3	1262	1	W13505	B. bronchiseptica	243	1.2	21	1	R07063	Human PsII2 expres
171	7	1.3	1263	1	W13503	B. pertussis adeny	244	1.2	21	1	P71703	Myc-his tag peptid
172	7	1.3	1445	1	R12108	N-terminal deleted	245	1.2	21	1	W48442	Human CS198 protei
173	7	1.3	1522	1	P93357	Sequence of the ca	246	1.2	21	1	W54281	Myc-His peptide ta
174	7	1.3	1644	1	W13504	B. bronchiseptica	247	1.2	21	1	W54435	Human CS198 protei
175	7	1.3	1645	1	W13502	B. pertussis adeny	248	1.2	21	1	W75875	Myc-his tag peptid
176	7	1.3	1705	1	P94365	Adeny cyclase fro	249	1.2	21	1	W79683	Human CS198 protei
177	7	1.3	1706	1	R08031	Adeny cyclase fro	250	1.2	21	1	W86030	Myc-his tag used 1
178	7	1.3	1891	1	W52844	Amycolatopsis medi	251	1.2	21	1	W95563	
179	7	1.3					252	1.2	21	1		

253	6	1.2	21	1	W92257	Myc-his tag sequen	326	6	1.2	94	1	R89747	AFT-1 interacting
254	6	1.2	21	1	W97661	Myc-his tag. Nucle	327	6	1.2	95	1	W77613	Deoxyribose-phosph
255	6	1.2	21	1	W95646	c-myc oncoprotein	328	6	1.2	97	1	W06408	Histidine tagged v
256	6	1.2	21	1	Y06793	Myc-His peptide ta	329	6	1.2	97	1	W06410	Histidine tagged v
257	6	1.2	21	1	Y07741	Human breast-speci	330	6	1.2	97	1	W06454	Recombinant human
258	6	1.2	22	1	W67267	Peptide #14 with c	331	6	1.2	99	1	W68490	Light chain of the
259	6	1.2	23	1	R79201	TGF-beta4 residues	332	6	1.2	100	1	W39892	Recombinant human
260	6	1.2	23	1	R94313	Histidine-6 insert	333	6	1.2	104	1	W89951	Antigen 1 from clu
261	6	1.2	23	1	W95472	Plasmid construct	334	6	1.2	105	1	W00920	Haspnp polypeptide
262	6	1.2	25	1	R79206	TGF-beta4 residues	335	6	1.2	105	1	Y04772	Mycobacterium spec
263	6	1.2	25	1	W10250	Hepatitis B virus	336	6	1.2	107	1	W75892	Peptide inhibitor
264	6	1.2	25	1	W67272	Peptide #19 with c	337	6	1.2	107	1	Y04785	Mycobacterium spec
265	6	1.2	26	1	R64926	Four Corners Virus	338	6	1.2	108	1	R51502	Pilin protein vari
266	6	1.2	26	1	W57233	Myelin basic prote	339	6	1.2	108	1	R51503	Pilin protein vari
267	6	1.2	28	1	W80575	N-terminal sequenc	340	6	1.2	108	1	Y04784	Mycobacterium spec
268	6	1.2	29	1	R66180	Portion of soluble	341	6	1.2	109	1	W39800	Variable domain of
269	6	1.2	30	1	W24708	Finger 2 domain of	342	6	1.2	109	1	Y00106	Enterococcus faeca
270	6	1.2	30	1	W65472	Hepatitis B surfac	343	6	1.2	112	1	R04027	X gene product of
271	6	1.2	31	1	W71427	Peptide having an	344	6	1.2	112	1	R34281	Human heavy chain
272	6	1.2	32	1	R97388	CC49 VH-spacer-TN	345	6	1.2	112	1	W16469	Tyca peptide/cy car
273	6	1.2	32	1	W71428	Peptide having an	346	6	1.2	113	1	W38475	S. pneumoniae. Nov
274	6	1.2	33	1	W71429	Peptide having an	347	6	1.2	116	1	W53994	Anti-CD4 antibody
275	6	1.2	33	1	W71445	Peptide having an	348	6	1.2	118	1	W08176	TGF-beta fusion pr
276	6	1.2	34	1	R37710	Delta14 Ser17 N-te	349	6	1.2	120	1	R42205	Valpha TCR gene w
277	6	1.2	34	1	W71430	Peptide having an	350	6	1.2	121	1	R42206	Vbeta TCR gene w
278	6	1.2	35	1	W71452	Peptide having an	351	6	1.2	122	1	W08181	TGF-beta fusion pr
279	6	1.2	35	1	W71453	Peptide having an	352	6	1.2	124	1	W08177	TGF-beta fusion pr
280	6	1.2	35	1	W71431	Peptide having an	353	6	1.2	124	1	W71693	R21-Hpro-28 fusion
281	6	1.2	36	1	W61376	GroEL N-terminal h	354	6	1.2	125	1	R80333	Protein polymeri
282	6	1.2	36	1	W65068	E. coli GroEL N-te	355	6	1.2	126	1	R65292	HIV-1 fusion prote
283	6	1.2	36	1	W71432	Peptide having an	356	6	1.2	126	1	R90543	pUc4-5-CDK-BP clon
284	6	1.2	36	1	W71437	Peptide having an	357	6	1.2	128	1	W08188	TGF-beta fusion pr
285	6	1.2	36	1	W71438	Peptide having an	358	6	1.2	130	1	W02119	Vector expressing
286	6	1.2	37	1	W71433	Peptide having an	359	6	1.2	130	1	W02119	DETI-DEI2-spacer-e
287	6	1.2	38	1	W71434	PCANAB vector enc	360	6	1.2	130	1	W03981	TGF-beta fusion pr
288	6	1.2	40	1	W95473	Artificial protein	361	6	1.2	130	1	W08185	DETI-DEI2-spacer-e
289	6	1.2	41	1	W34543	scfv hexahistidine	362	6	1.2	130	1	W11285	DETI-DEI2-spacer-e
290	6	1.2	43	1	R51891	Vector pUC119PHIS-	363	6	1.2	131	1	W19623	Human src SH2 doma
291	6	1.2	43	1	W19897	PCANAB vector enc	364	6	1.2	131	1	W08182	TGF-beta fusion pr
292	6	1.2	44	1	W95474	PELB signal sequen	365	6	1.2	132	1	R65290	Rat beta amyloid f
293	6	1.2	44	1	R75720	Staphylococcal ent	366	6	1.2	132	1	R65291	Human beta amyloid
294	6	1.2	45	1	W04488	Staphylococcal ent	367	6	1.2	132	1	W08189	TGF-beta fusion pr
295	6	1.2	45	1	W73917	Modified hsp60 gen	368	6	1.2	132	1	W75070	Human secreted pro
296	6	1.2	48	1	Y006970	Enterococcus faeca	369	6	1.2	133	1	W03121	DETI-DEI2-spacer-e
297	6	1.2	48	1	Y00107	Human factor VII E	370	6	1.2	133	1	W03983	DETI-DEI2-spacer-e
298	6	1.2	52	1	W80576	Hepatitis G virus	371	6	1.2	133	1	W20466	H. pylori transpor
299	6	1.2	54	1	W89178	Plasmid pOR4585 1n	372	6	1.2	133	1	W24668	TGF-beta fusion pr
300	6	1.2	55	1	R87580	Hepatitis G virus	373	6	1.2	133	1	W08178	DETI-DEI2-spacer-e
301	6	1.2	55	1	W76092	HGV Q9 epitope pro	374	6	1.2	133	1	W11287	Human hcp SH2 doma
302	6	1.2	55	1	W80186	HGV epitope sequen	375	6	1.2	133	1	W19625	Sequence of plant
303	6	1.2	55	1	W89481	Hepatitis G virus	376	6	1.2	134	1	P81139	Brassica seed acyl
304	6	1.2	57	1	W92793	US856134 Seq ID 2	377	6	1.2	134	1	R34979	B. campestris ACP.
305	6	1.2	57	1	W19696	Vector pUC119MCH-e	378	6	1.2	134	1	R76695	DETI-DEI2-spacer-e
306	6	1.2	58	1	W99092	Oropouche NP prote	379	6	1.2	134	1	W03120	DETI-DEI2-spacer-e
307	6	1.2	65	1	W94263	H6FXN12 fusion pr	380	6	1.2	134	1	W03982	ACP protein encode
308	6	1.2	67	1	R74528	Peptide encoded by	381	6	1.2	134	1	W15416	TGF-beta fusion pr
309	6	1.2	67	1	W80711	S. pneumoniae prot	382	6	1.2	134	1	W08187	TGF-beta fusion pr
310	6	1.2	68	1	Y12424	Human 5' EST secre	383	6	1.2	134	1	W08190	TGF-beta fusion pr
311	6	1.2	69	1	W94771	H6FXtr1pB fusion p	384	6	1.2	134	1	W11286	DETI-DEI2-spacer-e
312	6	1.2	72	1	R88391	Measles virus F pr	385	6	1.2	134	1	W11289	DETI-DEI2-spacer-e
313	6	1.2	72	1	W71692	Human defensin pre	386	6	1.2	134	1	W19624	Human lck SH2 doma
314	6	1.2	73	1	W94270	H6FXtr1pA fusion p	387	6	1.2	134	1	W19628	Human lck SH2 doma
315	6	1.2	73	1	W81027	Helix-loop-helix d	388	6	1.2	134	1	W30671	Brassica campestr
316	6	1.2	76	1	R13758	Beta gal IGF-1 fus	389	6	1.2	134	1	W30527	Brassica campestr
317	6	1.2	80	1	Y11514	S. pneumoniae prot	390	6	1.2	135	1	W08183	TGF-beta fusion pr
318	6	1.2	84	1	Y04775	Mycobacterium spec	391	6	1.2	136	1	W20850	H. pylori transpor
319	6	1.2	85	1	W69337	Mouse dectin-2 His	392	6	1.2	136	1	W08186	TGF-beta fusion pr
320	6	1.2	87	1	R82953	Multi-domain prote	393	6	1.2	137	1	P70101	Sequence encoded b
321	6	1.2	89	1	W19974	Fibronectin fragme	394	6	1.2	137	1	W08180	TGF-beta fusion pr
322	6	1.2	89	1	W23133	Human prostate pro	395	6	1.2	137	1	W08184	TGF-beta fusion pr
323	6	1.2	89	1	W57195	Fibronectin III-Q	396	6	1.2	137	1	W97107	Thermostable facto
324	6	1.2	89	1	W82985	Human fibronectin	397	6	1.2	139	1	W08179	TGF-beta fusion pr
325	6	1.2	92	1	W00944	CMV500-CREB2ZIP. N	398	6	1.2	139	1	W31556	Fibronectin-bindin

399	6	1.2	143	1	R54047	Sequence of the VR
400	6	1.2	143	1	Y04827	Mycobacterium spec
401	6	1.2	144	1	R15178	hcg histidine subs
402	6	1.2	145	1	W11288	DET1-DET2-spacer-e
403	6	1.2	145	1	W19626	Human Stat6 SH2 do
404	6	1.2	145	1	W94266	H6XtRIPB-UB fusio
405	6	1.2	146	1	W42691	Recombinant rabbit
406	6	1.2	147	1	R27495	Fusion protein exp
407	6	1.2	152	1	W94265	H6FXTN3 fusion pro
408	6	1.2	157	1	W01774	1-153 (delta 112-1
409	6	1.2	158	1	W01774	(His)6/1-153 (delt
410	6	1.2	159	1	W31552	Collagen binding P
411	6	1.2	159	1	W69170	N-terminally tagge
412	6	1.2	162	1	W42692	Recombinant human
413	6	1.2	165	1	R42838	Prochymosin N-term
414	6	1.2	166	1	W59153	L. lactis rhb C-t
415	6	1.2	166	1	W59162	L. lactis MG1316 r
416	6	1.2	171	1	R90765	CROC-1 c-fos promo
417	6	1.2	171	1	W48992	Amino acid sequenc
418	6	1.2	172	1	R37562	Human papilloma vi
419	6	1.2	172	1	R37563	Human papilloma vi
420	6	1.2	172	1	W26532	Trypanosoma cruzi
421	6	1.2	173	1	W60513	Antigenic determin
422	6	1.2	174	1	W06101	Proteinlipid protel
423	6	1.2	174	1	W95083	S. aureus phosphor
424	6	1.2	174	1	W95084	Phosphoribosyl tra
425	6	1.2	177	1	W68589	Staphylococcus aur
426	6	1.2	178	1	R95259	ABC protease P II
427	6	1.2	180	1	W69236	Mouse dectin-1 His
428	6	1.2	180	1	W94264	H6FXTN23 fusion pr
429	6	1.2	182	1	W72889	Mycobacterium tube
430	6	1.2	183	1	P60623	Sequence A encoded
431	6	1.2	185	1	W71501	Helicobacter poly
432	6	1.2	192	1	R80345	Protein polymeric
433	6	1.2	195	1	R97229	Recombinant Brucel
434	6	1.2	196	1	R75335	Human oncostatin-M
435	6	1.2	197	1	W70495	Human oncostatin-M
436	6	1.2	197	1	W94262	Human disease rela
437	6	1.2	198	1	R80346	H6FXTN123 fusion P
438	6	1.2	198	1	W57672	Protein polymeric
439	6	1.2	198	1	W49722	Collagen-like poly
440	6	1.2	201	1	W04573	Protein polymer ad
441	6	1.2	201	1	W12966	HCV NS3 construct
442	6	1.2	201	1	W01644	HCV NS3 soluble pr
443	6	1.2	203	1	W06107	Human NS3 soluble pr
444	6	1.2	203	1	W06108	Human NS3 soluble pr
445	6	1.2	204	1	R07470	Human NS3 soluble pr
446	6	1.2	204	1	W14318	Transcript of plas
447	6	1.2	206	1	R33381	Cytokine SOSM. Hyb
448	6	1.2	206	1	W70222	Leishmania antigen
449	6	1.2	210	1	W04582	HCV NS3 protease c
450	6	1.2	210	1	W09241	HCV NS3 protease c
451	6	1.2	210	1	W01650	HCV NS3 protease c
452	6	1.2	211	1	R06501	GST-47 encoded pro
453	6	1.2	211	1	W31553	Collagen binding P
454	6	1.2	214	1	R04472	Human Ras-like pro
455	6	1.2	217	1	R80347	Protein polymeric
456	6	1.2	218	1	W37339	Mouse HPRT. Multiv
457	6	1.2	218	1	W47019	Arabidopsis thalia
458	6	1.2	218	1	W47022	Arabidopsis thalia
459	6	1.2	218	1	W47018	Arabidopsis thalia
460	6	1.2	219	1	W23459	Recombinant MAM su
461	6	1.2	220	1	W47020	Arabidopsis thalia
462	6	1.2	223	1	R12899	Oncostatin M delet
463	6	1.2	223	1	R80348	Protein polymeric
464	6	1.2	223	1	R87020	BZLF2 protein. Eps
465	6	1.2	223	1	W47351	Epstein-Barr virus
466	6	1.2	224	1	R12898	Oncostatin M delet
467	6	1.2	224	1	R55278	HIV2(gp32)-CD4 ex
468	6	1.2	226	1	W18054	Recombinant human
469	6	1.2	226	1	W72902	Mycobacterium tube
470	6	1.2	226	1	Y03158	Recombinant human
471	6	1.2	226	1	Y03174	Recombinant human
472	6	1.2	227	1	P82309	Oncostatin M New D
473	6	1.2	227	1	R12871	Oncostatin M delet
474	6	1.2	227	1	R12896	Oncostatin M subtl
475	6	1.2	227	1	R12895	Oncostatin M subtl
476	6	1.2	227	1	R12897	Oncostatin M subtl
477	6	1.2	227	1	R12900	Oncostatin M subtl
478	6	1.2	227	1	R12901	Oncostatin M delet
479	6	1.2	227	1	W05530	Mouse TRADD intrac
480	6	1.2	228	1	W94261	CIH6FXTN123 fusio
481	6	1.2	230	1	R13593	Oncostatin M inser
482	6	1.2	233	1	R13203	Oncostatin M inser
483	6	1.2	233	1	R43011	Staphylococcal ent
484	6	1.2	233	1	W07613	Staphylococcal ent
485	6	1.2	233	1	W06738	Staphylococcal ent
486	6	1.2	233	1	W35373	Staphylococcal ent
487	6	1.2	233	1	W45516	Staphylococcal ent
488	6	1.2	233	1	W64670	NANC-2 heavy chal
489	6	1.2	234	1	W69607	Human UC PANCA mon
490	6	1.2	234	1	W69608	Human cysteine mod
491	6	1.2	235	1	R80349	Human cysteine mod
492	6	1.2	235	1	W69611	Protein polymeric
493	6	1.2	236	1	W69612	Human cysteine mod
494	6	1.2	236	1	W37700	Human cysteine mod
495	6	1.2	242	1	P70483	Protein sequence o
496	6	1.2	243	1	W37701	Protein sequence o
497	6	1.2	244	1	W07614	Protein sequence o
498	6	1.2	244	1	W64672	Ulcerative colitis
499	6	1.2	244	1	W90309	Human UC PANCA mon
500	6	1.2	245	1	W35375	Human UC PANCA mon
501	6	1.2	246	1	W37703	Staphylococcus ent
502	6	1.2	247	1	R42382	Protein sequence o
503	6	1.2	248	1	R42208	H. somnus lipa. Ha
504	6	1.2	248	1	W63830	scvAlphabetaCDP1b
505	6	1.2	250	1	R42307	Transformant CDM/1
506	6	1.2	250	1	W09239	scvAlphabeta with
507	6	1.2	250	1	W01648	HCV NS3 soluble pr
508	6	1.2	250	1	W14944	Chimeric human/por
509	6	1.2	251	1	W49735	Protein polymer ad
510	6	1.2	252	1	P92082	Protein polymeric
511	6	1.2	252	1	R33380	Consensus sequence
512	6	1.2	252	1	R88202	Cytokine hOSM. Hyb
513	6	1.2	253	1	P94685	Human oncostatin M
514	6	1.2	253	1	W80667	Amino acid sequenc
515	6	1.2	253	1	W90225	S. pneumoniae lacra
516	6	1.2	254	1	W41995	Anti-B7.1 monosp
517	6	1.2	254	1	W41997	Anti-B7.1 monosp
518	6	1.2	254	1	W41997	Flea cysteine prot
519	6	1.2	255	1	W09237	HCV NS3-NS4A prote
520	6	1.2	255	1	W01646	HCV NS3 soluble pr
521	6	1.2	255	1	W37702	Protein sequence o
522	6	1.2	255	1	W47147	Chloramphenicol ac
523	6	1.2	256	1	W54875	Staphylococcus aur
524	6	1.2	258	1	W90221	Anti-B7.1 monosp
525	6	1.2	260	1	R77617	Anti-B7.1 monosp
526	6	1.2	261	1	R82955	Anti-C5 Mab N19/8
527	6	1.2	262	1	R97726	Three domain prote
528	6	1.2	262	1	W49734	B10 single chain T
529	6	1.2	263	1	P93565	Protein polymer ad
530	6	1.2	263	1	P90507	Sequence of an epi
531	6	1.2	263	1	R82957	Sequence of an epi
532	6	1.2	263	1	R97725	Three domain prote
533	6	1.2	263	1	R95968	D10 single chain T
534	6	1.2	263	1	R95969	Nucleic acid recog
535	6	1.2	263	1	R95969	Nucleic acid recog
536	6	1.2	263	1	W32447	Mycobacterium tube
537	6	1.2	263	1	W32379	Mycobacterium tube
538	6	1.2	263	1	W64317	Mycobacterium tube
539	6	1.2	263	1	W81680	Myobacterium tube
540	6	1.2	263	1	W90226	Anti-B7.2 monosp
541	6	1.2	264	1	R81428	Hepatitis GB virus
542	6	1.2	266	1	P91236	(Hexahis) (mdhFR)E
543	6	1.2	268	1	R97561	Human papilloma vi
544	6	1.2	268	1	W61207	Streptococcus pneu
545	6	1.2	268	1	W90222	Anti-B7.2 monosp

545	6	1.2	269	1	R40801	618	6	1.2	343	1	R82455	Hepatitis GB virus
546	6	1.2	270	1	W04571	619	6	1.2	343	1	W60053	Human protein tyro
547	6	1.2	270	1	W09236	620	6	1.2	344	1	W08097	Vmpt soluble varia
548	6	1.2	270	1	W01645	621	6	1.2	344	1	W18879	Leucine aminopepti
549	6	1.2	272	1	W26698	622	6	1.2	345	1	W15555	Fibronectin-bindin
550	6	1.2	272	1	W87637	623	6	1.2	346	1	R39493	Human apoAIV mutel
551	6	1.2	274	1	W39899	624	6	1.2	348	1	P70593	Fructose-bisphosph
552	6	1.2	275	1	R84563	625	6	1.2	349	1	R43911	Sequence encoded b
553	6	1.2	276	1	W90227	626	6	1.2	349	1	W06182	Modified HCPB (D25
554	6	1.2	277	1	P70298	627	6	1.2	349	1	W06173	mature HCPB-(H195
555	6	1.2	277	1	W33275	628	6	1.2	349	1	W06181	Modified HCPB (H195
556	6	1.2	281	1	W49736	629	6	1.2	349	1	W13750	Carboxypeptidase B
557	6	1.2	281	1	W82744	630	6	1.2	349	1	W13751	Carboxypeptidase B
558	6	1.2	281	1	W90223	631	6	1.2	349	1	W13752	Carboxypeptidase B
559	6	1.2	282	1	Y13365	632	6	1.2	349	1	W13753	Carboxypeptidase B
560	6	1.2	284	1	W64366	633	6	1.2	349	1	W13757	Carboxypeptidase B
561	6	1.2	284	1	W81733	634	6	1.2	349	1	W13758	Carboxypeptidase B
562	6	1.2	286	1	R49836	635	6	1.2	349	1	W13759	Carboxypeptidase B
563	6	1.2	286	1	W98520	636	6	1.2	349	1	W13760	Carboxypeptidase B
564	6	1.2	287	1	W49737	637	6	1.2	349	1	W13761	Carboxypeptidase B
565	6	1.2	288	1	W82316	638	6	1.2	349	1	W13762	Carboxypeptidase B
566	6	1.2	288	1	W82317	639	6	1.2	349	1	W13763	Carboxypeptidase B
567	6	1.2	288	1	W82482	640	6	1.2	349	1	W13764	Carboxypeptidase B
568	6	1.2	288	1	W82743	641	6	1.2	349	1	W13765	Carboxypeptidase B
569	6	1.2	291	1	W49738	642	6	1.2	349	1	W13749	Carboxypeptidase B
570	6	1.2	297	1	R56543	643	6	1.2	351	1	W56000	Fibroblast growth
571	6	1.2	298	1	W15008	644	6	1.2	353	1	R81355	Thrombopoietin pro
572	6	1.2	301	1	W15007	645	6	1.2	353	1	R81356	Thrombopoietin pro
573	6	1.2	301	1	W73217	646	6	1.2	355	1	Y06997	Galectin-9 protein
574	6	1.2	303	1	W08956	647	6	1.2	356	1	W22676	Extracellular doma
575	6	1.2	304	1	R73599	648	6	1.2	357	1	R66810	His-tagged polylin
576	6	1.2	308	1	W13322	649	6	1.2	359	1	W61046	Human GSBP2. New C
577	6	1.2	308	1	W34322	650	6	1.2	360	1	R75006	Tomato S-adenosyl-
578	6	1.2	309	1	W83322	651	6	1.2	360	1	W37908	Tbv25-Pfs28C fusio
579	6	1.2	310	1	W83324	652	6	1.2	360	1	W37908	Cytokine suppress
580	6	1.2	311	1	W56504	653	6	1.2	360	1	W49066	Human monocytic CSA
581	6	1.2	312	1	W05528	654	6	1.2	360	1	W37908	Tbv25-28 fusion pr
582	6	1.2	312	1	W83323	655	6	1.2	363	1	W37910	Prothymosin. DNA t
583	6	1.2	315	1	R96268	656	6	1.2	365	1	P94144	Sequence of calf p
584	6	1.2	316	1	Y01820	657	6	1.2	365	1	R05080	Sequence encoded b
585	6	1.2	317	1	W61195	658	6	1.2	365	1	P30603	Case-A2 VI/V2 fusi
586	6	1.2	319	1	W35855	659	6	1.2	368	1	W01808	Human papillomavir
587	6	1.2	320	1	W24210	660	6	1.2	368	1	W01808	H. pylori GHPo 202
588	6	1.2	320	1	W38217	661	6	1.2	369	1	W98612	Renal cancer assoc
589	6	1.2	320	1	W59937	662	6	1.2	372	1	W07080	Mammalian Zap-70 p
590	6	1.2	320	1	W79965	663	6	1.2	372	1	W82481	Human apoAIV mutel
591	6	1.2	323	1	W06407	664	6	1.2	373	1	R39486	MP4 chimera (MBP21
592	6	1.2	326	1	W06409	665	6	1.2	373	1	W06103	Mycobacterium spec
593	6	1.2	329	1	W44778	666	6	1.2	373	1	Y07203	Murine promiscuous
594	6	1.2	330	1	W94267	667	6	1.2	374	1	W42434	Sequence encoded b
595	6	1.2	331	1	P70328	668	6	1.2	375	1	P40078	PM4 chimera (delta
596	6	1.2	331	1	W94268	669	6	1.2	375	1	W06104	Human papillomavir
597	6	1.2	332	1	W14919	670	6	1.2	375	1	W01809	Carbolic anhydrase
598	6	1.2	332	1	W32418	671	6	1.2	376	1	R21416	Chlamydomonas carb
599	6	1.2	332	1	W32350	672	6	1.2	377	1	R21417	Aminopeptidase. En
600	6	1.2	332	1	W64322	673	6	1.2	377	1	W05589	Human fc receptor
601	6	1.2	332	1	W76189	674	6	1.2	377	1	W86195	NANB virus antige
602	6	1.2	332	1	W59951	675	6	1.2	378	1	R50076	SELPOR polymer. T1
603	6	1.2	333	1	W81683	676	6	1.2	378	1	W09219	Amino acid sequenc
604	6	1.2	333	1	R31019	677	6	1.2	378	1	W53545	Multiple sclerosis
605	6	1.2	334	1	W98714	678	6	1.2	378	1	W71072	Protein encoded by
606	6	1.2	337	1	P81191	679	6	1.2	378	1	W95558	Sequence of an epi
607	6	1.2	337	1	R39485	680	6	1.2	379	1	P90508	Sequence of an epi
608	6	1.2	338	1	R80350	681	6	1.2	379	1	P30013	Sequence of an epi
609	6	1.2	339	1	W37909	682	6	1.2	380	1	P90510	Sequence of an epi
610	6	1.2	341	1	W37907	683	6	1.2	380	1	P90511	Sequence of an epi
611	6	1.2	342	1	R39487	684	6	1.2	380	1	P90512	Sequence of new po
612	6	1.2	342	1	R39489	685	6	1.2	380	1	P20038	Pre-prorenin-A pr
613	6	1.2	342	1	R39491	686	6	1.2	380	1	R37714	hCNTF fusion proti
614	6	1.2	342	1	R39496	687	6	1.2	380	1	W19111	Human tumour succe
615	6	1.2	342	1	R39498	688	6	1.2	381	1	R20730	Prochymosin (prote
616	6	1.2	342	1	R95054	689	6	1.2	381	1	P40218	Sequence of rennin
617	6	1.2	342	1	W55958	690	6	1.2	381	1		

691	6	1.2	381	1	P40559	Sequence of a poly	764	6	1.2	474	1	W36132	N-terminal modifie
692	6	1.2	381	1	P30086	Sequence encoded b	765	6	1.2	477	1	R43399	Human OTC. Gene mu
693	6	1.2	381	1	P30446	Sequence encoded b	766	6	1.2	479	1	W49739	Protein polymer ad
694	6	1.2	381	1	W19110	Mouse tumour susce	767	6	1.2	482	1	Y06915	Amino acid sequenc
695	6	1.2	381	1	W93425	Mouse tsg101 prote	768	6	1.2	491	1	W13373	Recombinant Sp 100
696	6	1.2	383	1	W48336	Brucella abortus 3	769	6	1.2	492	1	W06105	MMOGp4 chimera (MB
697	6	1.2	386	1	W41786	Cyclin E (MRHHNH	770	6	1.2	493	1	R13392	p. falciparum spor
698	6	1.2	386	1	W37242	Globodera rostocki	771	6	1.2	494	1	R90765	Human K+ channel 2
699	6	1.2	386	1	W92792	Human TNF receptor	772	6	1.2	494	1	W42996	Putative mature po
700	6	1.2	386	1	W99018	Human TNF receptor	773	6	1.2	495	1	W37062	HIV-1 breakthrough
701	6	1.2	386	1	W99019	Human TRAIL recept	774	6	1.2	495	1	W37063	HIV-1 breakthrough
702	6	1.2	386	1	Y04144	Human TANGO-74 pro	775	6	1.2	496	1	R95058	GALA-DT-IL-2 multi
703	6	1.2	386	1	W98200	RTD, inhibitor of	776	6	1.2	498	1	R28805	Soluble HIV gp120.
704	6	1.2	390	1	P94663	New sarcosine oxid	777	6	1.2	511	1	W78475	Autographa callifor
705	6	1.2	390	1	W07379	Bacillus sarcosine	778	6	1.2	512	1	W31554	Collagen binding p
706	6	1.2	390	1	W93424	Human NHTS protein	779	6	1.2	514	1	W31876	GFP variants S65C
707	6	1.2	391	1	R20011	Soybean seed stear	780	6	1.2	514	1	W83359	Streptococcus pneu
708	6	1.2	391	1	R82024	Stearoyl-ACP-desat	781	6	1.2	514	1	W88426	Chlamydia pneumoni
709	6	1.2	391	1	R97659	Single chain T cel	782	6	1.2	515	1	P91235	(ENV-80)(GAG-VII)
710	6	1.2	391	1	W32449	Mycobacterium tube	783	6	1.2	519	1	W00181	HIV-1 env protein.
711	6	1.2	391	1	W32381	Mycobacterium tube	784	6	1.2	521	1	R69997	HIV-1 env protein.
712	6	1.2	391	1	W37939	Amino acid sequenc	785	6	1.2	524	1	R56525	Protein kinase (NU
713	6	1.2	391	1	W64335	Mycobacterium tube	786	6	1.2	524	1	R76621	Saccharomyces cere
714	6	1.2	391	1	W81702	M. tuberculosis Im	787	6	1.2	525	1	R99466	Biosynthetic enzym
715	6	1.2	391	1	Y04778	Mycobacterium spec	788	6	1.2	527	1	R82954	Six domain protein
716	6	1.2	392	1	Y00139	Enterococcus faeca	789	6	1.2	530	1	R95053	scFv(FRP5)-DETA-DG
717	6	1.2	394	1	Y04779	Mycobacterium spec	790	6	1.2	530	1	W28501	Birch pollen co-fa
718	6	1.2	395	1	R88225	Transcription fact	791	6	1.2	530	1	W47126	Uridine diphospho-
719	6	1.2	397	1	W80573	Human O-fucosyltra	792	6	1.2	530	1	W78476	Baculovirus ISF pr
720	6	1.2	397	1	W81598	Protein encoded by	793	6	1.2	532	1	R42457	Enzyme involved in
721	6	1.2	400	1	P91941	Sequence of prepro	794	6	1.2	532	1	R83016	Recombinant papill
722	6	1.2	400	1	W55691	H. pylori ORF 069p	795	6	1.2	533	1	W19763	p53-GM-CSF immuno
723	6	1.2	406	1	R53528	Thyroid hormone re	796	6	1.2	537	1	Y03839	Chimeric transcrip
724	6	1.2	406	1	W06100	Human 26S proteaso	797	6	1.2	539	1	W48358	Fractionated X irr
725	6	1.2	406	1	W85089	Thyroid hormone re	798	6	1.2	544	1	W46314	AAV4 VP3 coat prot
726	6	1.2	406	1	W92374	Human JLI protein.	799	6	1.2	546	1	W38234	VEGF/CPG2 fusion p
727	6	1.2	407	1	R88755	Flock house virus	800	6	1.2	547	1	W22707	N-terminal truncat
728	6	1.2	407	1	R88758	FHV capsid protein	801	6	1.2	550	1	R42085	NF-KappaB p65 prot
729	6	1.2	409	1	R88756	FHV capsid protein	802	6	1.2	550	1	R42254	Human p65 protein
730	6	1.2	409	1	W43449	Tobacco laccase cl	803	6	1.2	550	1	R42255	Human p65 protein
731	6	1.2	410	1	R88757	FHV capsid protein	804	6	1.2	551	1	W38238	VEGF/CPG2 fusion p
732	6	1.2	415	1	R88759	FHV capsid protein	805	6	1.2	552	1	W48873	Hypophyza sp. str
733	6	1.2	418	1	W97650	Soybean SUG1 polyp	806	6	1.2	552	1	W96299	Vgag streptogramin
734	6	1.2	421	1	R32712	Rat intrinsic fact	807	6	1.2	553	1	W22706	Arginyl-tRNA synth
735	6	1.2	421	1	R32733	Rat intrinsic fact	808	6	1.2	553	1	W11508	Single chain anti-
736	6	1.2	423	1	R95055	IL-2-DETA-DGALA mu	809	6	1.2	553	1	W73223	H22-anti-CEA antib
737	6	1.2	423	1	W01513	An Escherichia col	810	6	1.2	554	1	R70827	Anti-catact immu
738	6	1.2	425	1	R60698	Fragment of the hu	811	6	1.2	554	1	W56099	A. oryzae ATCC2038
739	6	1.2	432	1	W38241	Carboxypeptidase C	812	6	1.2	556	1	W98849	H. pylori GHP 160
740	6	1.2	436	1	R94686	Threonine dehydrat	813	6	1.2	556	1	W90218	Bispecific tetrava
741	6	1.2	436	1	R94687	Threonine dehydrat	814	6	1.2	559	1	W78753	Human EXAI-A. Nucl
742	6	1.2	436	1	R94688	Threonine dehydrat	815	6	1.2	561	1	R48058	Sequence of protea
743	6	1.2	436	1	R94689	Threonine dehydrat	816	6	1.2	564	1	W48788	Thyroid peroxidase
744	6	1.2	436	1	R94690	Threonine dehydrat	817	6	1.2	564	1	P82595	Amylase from strep
745	6	1.2	436	1	R94691	Threonine dehydrat	818	6	1.2	566	1	R08263	Myobacterium spec
746	6	1.2	437	1	W37738	Nucleotide sequenc	819	6	1.2	572	1	Y04954	Hypophyza sp. str
747	6	1.2	441	1	W05410	Mouse H74 protein.	820	6	1.2	573	1	W48874	Delta-endotoxin pr
748	6	1.2	443	1	W94460	NADH:H2O oxidase a	821	6	1.2	575	1	R11862	Bispecific tetrava
749	6	1.2	445	1	W68391	Clostridium botuli	822	6	1.2	580	1	W90217	Human ifizzled-5 p
750	6	1.2	445	1	W97709	Staphylococcus aur	823	6	1.2	585	1	W31871	Recombinant furin
751	6	1.2	446	1	Y04126	Bacterial and mamm	824	6	1.2	591	1	W21963	VEGF/CPG2 fusion p
752	6	1.2	450	1	P94370	Sequence encoded b	825	6	1.2	592	1	W38233	Recombinant furin
753	6	1.2	450	1	W46606	Tyrosine kinase as	826	6	1.2	592	1	W78754	Human EXAI-B. Nucl
754	6	1.2	453	1	R63790	Aspergillus aculea	827	6	1.2	592	1	W94269	H6F5SCFV(CRA6)trip
755	6	1.2	457	1	W25789	Acromonium cellulo	828	6	1.2	593	1	R86292	Pseudomonas mesoac
756	6	1.2	458	1	P94376	BamHI/SalI insert	829	6	1.2	596	1	W48785	Thyroid peroxidase
757	6	1.2	459	1	R98905	Vector SLC1-derivate	830	6	1.2	597	1	W21964	Recombinant furin
758	6	1.2	459	1	W29212	SSCI single chain	831	6	1.2	598	1	W46313	Human papillomavir
759	6	1.2	463	1	W05137	TGF alpha-ETA fusi	832	6	1.2	601	1	W03562	Human papillomavir
760	6	1.2	465	1	W01810	Human papillomavir	833	6	1.2	601	1	W44291	Human papillomavir
761	6	1.2	469	1	W68511	B. flavum fumase	834	6	1.2	604	1	P60359	Glucanase struc
762	6	1.2	472	1	W37243	Globodera rostocki	835	6	1.2	606	1	W85019	p38-green fluoresc
763	6	1.2	472	1	W86337	Kidney injury asso	836	6	1.2	607	1	W38237	VEGF/CPG2 fusion p

883 1.2 942 1 W56699
 984 1.2 948 1 W48781
 985 1.2 956 1 R15784
 986 1.2 966 1 W09223
 987 1.2 970 1 R85089
 988 1.2 974 1 W64221
 989 1.2 976 1 W02289
 990 1.2 980 1 R32359
 991 1.2 985 1 R42995
 992 1.2 985 1 R42214
 993 1.2 986 1 W13670
 994 1.2 990 1 R39343
 995 1.2 994 1 R87018
 996 1.2 994 1 W26366
 997 1.2 995 1 R75712
 998 1.2 995 1 R84658
 999 1.2 1002 1 W09218
 1000 1.2 1002 1 W53544

ALIGNMENTS

Modified protein K
 Thyroid peroxidase
 B.thuringiensis to
 SERP-CSI2 polymer
 EPR-like receptor
 Human secreted pro
 Mouse neuron resti
 Human KA-2 recepto
 Glycosyltransferas
 Aspergillus niger
 C-proteinase encod
 Epib protein. Nove
 Receptor. tyrosine
 Mouse Nuk tyrosine
 Epi-related PTK Ce
 Strivan-p pol prot
 SERP-CSI2 polymer. T1
 Amino acid sequenc

RESULT 1
 ID R79961 standard; Protein; 329 AA.
 AC R79961;
 DT 18-APR-1996 (first entry)
 DE Partial lcrv (V antigen) of Y. pestis.
 KW lcrv; V antigen; virulence; plaque; vaccine; epitope.
 OS Yersinia pestis.
 PN W09524475-A1.
 PD 14-SEP-1995.
 PR 06-MAR-1995; G00481.
 PR 08-MAR-1994; GB-004577.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Leary SEC, Titball RW, Williamson ED, Leary SE;
 DR WPI; 95-328268/42.
 DR N-PSDB; T04222.
 DT Recombinant DNA expressing Yersinia pestis V antigen - useful in
 PI oral or parenteral vaccines for protection against plague
 PS Claim 6: page 11-13; 25pp, English.
 CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrv) encoding
 CC all or a protective epitopic part of the mature V protein of Yersinia
 CC pestis. The protein was expressed as a fusion protein with maltose
 CC binding protein or glutathione-S-transferase in 3 different plasmid
 CC vectors. Y. pestis is the highly virulent causative organism of plague
 CC in a wide range of animals, including man. The V antigen (lcrv) is an
 CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
 CC The V antigen is postulated to act as a virulence antigen, and
 CC transformed microorganisms contg. recombinant DNA encoding a V antigen
 CC protein/peptide are useful in vaccines to protect against plague.
 SQ Sequence 329 AA;

Query Match 46.6%; Score 243; DB 1; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.2e-225;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHYDNOQNGIKRVEEFLSSPNTQWELRAFMVHFSITADRIDDILKIYVDSM 338
 DB 87 ILKGGHYDNOQNGIKRVEEFLSSPNTQWELRAFMVHFSITADRIDDILKIYVDSM 146
 QY 339 NHGDAKSKLREELAEIETAEIKIYVIAQAEINKHLSSGTTINHDKSLNMDKNLYGYTD 398
 DB 147 NHGDAKSKLREELAEIETAEIKIYVIAQAEINKHLSSGTTINHDKSLNMDKNLYGYTD 206
 QY 399 EEIFKASAEYKILKEMPTTIOVDGSEKKIYVINDFLGSEKRRGTGALGNLKNSSYKNKN 458
 DB 207 EEIFKASAEYKILKEMPTTIOVDGSEKKIYVINDFLGSEKRRGTGALGNLKNSSYKNKN 266
 QY 459 NELSHFATTCSDKSRPLNDLVSOQKTQSDITSRFSNSAIEALNRFIOKYDSVMORLLDDT 518

DB 267 NELSHFATTCSDKSRPLNDLVSOQKTQSDITSRFSNSAIEALNRFIOKYDSVMORLLDDT 326
 QY 519 SGK 521
 DB 327 SGK 329

RESULT 2
 ID W01040 standard; Protein; 329 AA.
 AC W01040;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis V antigen.
 KW Plaque; vaccine; genetic immunisation; V antigen; lcrv;
 OS Yersinia pestis strain GB.
 FH Key location/Qualifiers
 FT peptide
 FT 1.4
 PN W09628551-A1.
 PD 19-SEP-1996.
 PR 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
 DR WPI; 96-433824/43.
 DR N-PSDB; T38242.
 DT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Disclosure: page 25-28; 98pp; English.
 CC Yersinia pestis V antigen (W01040) is capable of evoking protective
 CC immune responses in animals. A combined vaccine of V and F1
 CC antigens (see also W01042) can at least match the protection
 CC afforded by live attenuated EV76 vaccine without any of the hazards
 CC that have kept the EV vaccine from general use. The V antigen is
 CC produced using the lcrv gene (see also T38242) obtd. from Y.
 CC pestis by PCR amplification. It can also be prepd. as a fusion with
 CC F1 antigen (see also W01044-45) and expressed by attenuated Arco or
 CC C Salmonella typhi as a live vaccine for long-term protection
 CC against plague. Expression by gut-colonising bacterial transformants
 CC produces a protective response against Y. pestis.
 SQ Sequence 329 AA;

Query Match 46.6%; Score 243; DB 1; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.2e-225;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHYDNOQNGIKRVEEFLSSPNTQWELRAFMVHFSITADRIDDILKIYVDSM 338
 DB 87 ILKGGHYDNOQNGIKRVEEFLSSPNTQWELRAFMVHFSITADRIDDILKIYVDSM 146
 QY 339 NHGDAKSKLREELAEIETAEIKIYVIAQAEINKHLSSGTTINHDKSLNMDKNLYGYTD 398
 DB 147 NHGDAKSKLREELAEIETAEIKIYVIAQAEINKHLSSGTTINHDKSLNMDKNLYGYTD 206
 QY 399 EEIFKASAEYKILKEMPTTIOVDGSEKKIYVINDFLGSEKRRGTGALGNLKNSSYKNKN 458
 DB 207 EEIFKASAEYKILKEMPTTIOVDGSEKKIYVINDFLGSEKRRGTGALGNLKNSSYKNKN 266
 QY 459 NELSHFATTCSDKSRPLNDLVSOQKTQSDITSRFSNSAIEALNRFIOKYDSVMORLLDDT 518
 DB 267 NELSHFATTCSDKSRPLNDLVSOQKTQSDITSRFSNSAIEALNRFIOKYDSVMORLLDDT 326
 QY 519 SGK 521
 DB 327 SGK 329

RESULT 3
 W01041

ID W01041 standard; Protein; 329 AA.
AC W01041;
DT 28-DEC-1996 (first entry)
DE Y. pestis V antigen.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
OS Yersinia pestis strain GB.
FH Key
FT Location/Qualifiers
FT peptide 1..4
PT /note="vector-encoded peptide"
PN W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI; 96-433824/43.
DR N-PSDB; T38243.
PT Yersinia pestis V antigen and F1 antigen or their protective
PS epitopic parts - useful in vaccine for protection against plague
CC Disclosure; Page 32-35; 98pp; English.
CC Yersinia pestis V antigen (W01041) is capable of evoking protective
CC immune responses in animals. A combined vaccine of V and F1
CC antigens (see also W01042) can at least match the protection
CC afforded by live attenuated EV76 vaccine without any of the hazards
CC that have kept the lcrV gene (see also T38243) out of the V antigen is
CC produced using the lcrV gene (see also T38243) obtd. from Y. pestis
CC by PCR amplification. It can also be prepd. as a fusion with F1
CC antigen (see also W01044-45) and expressed by attenuated Arco or C
CC Salmonella typhimurium and gut-colonising bacteria for protection against
CC plague.
SQ Sequence 329 AA;

Query Match 46.6%; Score 243; DB 1; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.2e-225;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHNDLQNGIKRKEFELESSPNTQWELRAFMVMEFSLADRIDDDILKVIYDSM 338
DB 87 ILKGGHNDLQNGIKRKEFELESSPNTQWELRAFMVMEFSLADRIDDDILKVIYDSM 146
QY 339 NHHGDASKRLREELAEIETAEIKIYVIOAEINHLSSSGTINIHDKSINLMDKNLYGYTD 398
DB 147 NHHGDASKRLREELAEIETAEIKIYVIOAEINHLSSSGTINIHDKSINLMDKNLYGYTD 206
QY 399 EEIFKASAEYKILKEMQTTIOVDGSEKKIYVIOAEINHLSSSGTINIHDKSINLMDKNLYGYTD 458
DB 207 EEIFKASAEYKILKEMQTTIOVDGSEKKIYVIOAEINHLSSSGTINIHDKSINLMDKNLYGYTD 266
QY 459 NELSHFATTCSDKSRPLNDLVSOIKTQSLDITSRNSAIEALNRFIOKYYDSVMORLLDDT 518
DB 267 NELSHFATTCSDKSRPLNDLVSOIKTQSLDITSRNSAIEALNRFIOKYYDSVMORLLDDT 326
QY 519 SGK 521
DB 327 SGK 329

RESULT 4
ID W01045
AC W01045 standard; Protein; 480 AA.
DT 28-DEC-1996 (first entry)
DE Y. pestis F1/V antigen fusion.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
OS Yersinia pestis strain GB.
FH Key
FT Location/Qualifiers
FT protein 1..149

FT /label="F1 antigen"
FT /note="mature F1 antigen"
FT peptide 150..155
FT /label="linker"
FT /note="6-amino acid peptide linker"
FT protein 156..480
FT /label="V antigen"
FT /note="mature V antigen"

W09628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI; 96-433824/43.
DR N-PSDB; T38256.
PT Yersinia pestis V antigen and F1 antigen or their protective
PS epitopic parts - useful in vaccine for protection against plague
CC Disclosure; Page 51-55; 98pp; English.
CC A fusion protein (W01045) comprises the F1 antigen (see also
CC W01042) and V antigen (see also W01041) of Yersinia pestis joined
CC by a linker that allows each protein to attain its conformational
CC state. It is the product of a gene fusion (T38256) obtd. by PCR
CC amplification of Y. pestis DNA. F1/V fusion protein can be
CC expressed by gut-colonising organism transformants, to induce
CC an immune response against Y. pestis, the causative organism of
CC plague.
SQ Sequence 480 AA;

Query Match 39.7%; Score 207; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 6.7e-191;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGGHNDLQNGIKRKEFELESSPNTQWELRAFMVMEFSLADRIDDDILKVIYDSM 338
DB 238 ILKGGHNDLQNGIKRKEFELESSPNTQWELRAFMVMEFSLADRIDDDILKVIYDSM 297
QY 339 NHHGDASKRLREELAEIETAEIKIYVIOAEINHLSSSGTINIHDKSINLMDKNLYGYTD 398
DB 298 NHHGDASKRLREELAEIETAEIKIYVIOAEINHLSSSGTINIHDKSINLMDKNLYGYTD 357
QY 399 EEIFKASAEYKILKEMQTTIOVDGSEKKIYVIOAEINHLSSSGTINIHDKSINLMDKNLYGYTD 458
DB 358 EEIFKASAEYKILKEMQTTIOVDGSEKKIYVIOAEINHLSSSGTINIHDKSINLMDKNLYGYTD 417
QY 459 NELSHFATTCSDKSRPLNDLVSOIKTQ 485
DB 418 NELSHFATTCSDKSRPLNDLVSOIKTQ 444

RESULT 5
ID R76528
AC R76528 standard; Protein; 170 AA.
DT 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhimurium;
OS Yersinia pestis.
FH Key
FT Location/Qualifiers
FT protein 156..480
FT /label="V antigen"
FT /note="mature V antigen"

W09518231-A1.
PD 06-JUL-1995.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA) UK SEC FOR DEFENCE.
PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI; 95-246396/32.
DR N-PSDB; Q92819.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.

PS Disclosure: Page 20; 27pp; English.
 CC The sequence represents the Y. pestis cafi (F1) antigen expressed
 CC from plasmid pF0R1b. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated salmonella typhimurium or salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 170 AA;

Query Match 32.6%; Score 170; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1e-155;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 83
 DB 1 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 60

QY 84 ELLVGLTLGGYKTKTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 143
 DB 61 ELLVGLTLGGYKTKTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 120

QY 144 PKNVGENLVGDVVLATGSDQDFVRSISGKGKLAAGKXTDAVTVYSNQ 193
 DB 121 PKNVGENLVGDVVLATGSDQDFVRSISGKGKLAAGKXTDAVTVYSNQ 170

RESULT 6
 ID W01043 standard; Protein: 170 AA.
 AC W01043;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1 antigen (including signal peptide);
 KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;
 OS Yersinia pestis strain GB.
 FH Key Location/Qualifiers
 FT peptide 1..21
 FT /label= Sig_peptide
 FN W09628551-A1.
 PD 19-SEP-1996.
 PF 13-MAR-1996: G00571.
 PR 13-MAR-1995: GB-005059.
 PR 15-SEP-1995: GB-018946.
 PR 05-DEC-1995: GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
 DR N-PSDB; T38248.
 PT Yersinia pestis - useful in vaccine for protection against plague
 PT epitopic pests v antigen and F1 antigen of their protective
 PS Example 2; Page 61-62; 98pp; English.
 CC The F1 antigen (W01043), including the signal peptide, of Yersinia
 CC pestis was produced from a DNA sequence (T38248) obtd. by PCR
 CC amplification (see also T38257-58) of Y. pestis DNA. Expression
 CC of the F1 antigen (see also W01042) by gut-colonising organisms in
 CC the form of live vaccines can be used to protect an animal,
 CC including humans, against plague.
 SQ Sequence 170 AA;

Query Match 32.6%; Score 170; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1e-155;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 83
 DB 1 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 60

QY 84 ELLVGLTLGGYKTKTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 143

DB 61 ELLVGLTLGGYKTKTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 120

QY 144 PKNVGENLVGDVVLATGSDQDFVRSISGKGKLAAGKXTDAVTVYSNQ 193
 DB 121 PKNVGENLVGDVVLATGSDQDFVRSISGKGKLAAGKXTDAVTVYSNQ 170

RESULT 7
 ID W59782 standard; Protein: 170 AA.
 AC W59782;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYpF1sec170.
 KW F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-333331/29.
 DR N-PSDB; V41594.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 52; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 170 AA;

Query Match 32.6%; Score 170; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1e-155;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 83
 DB 1 MKRISVIAIALEFGTATANAADLTASTATATLVEPARITLTYKGAFTTMDNGNIDT 60

QY 84 ELLVGLTLGGYKTKTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 143
 DB 61 ELLVGLTLGGYKTKTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRDFDIS 120

QY 144 PKNVGENLVGDVVLATGSDQDFVRSISGKGKLAAGKXTDAVTVYSNQ 193
 DB 121 PKNVGENLVGDVVLATGSDQDFVRSISGKGKLAAGKXTDAVTVYSNQ 170

RESULT 8
 ID W59783 standard; Protein: 170 AA.
 AC W59783;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen nYpF1sec310.
 KW F1 antigen; plasmid; vaccine; plague; ds.
 OS Yersinia pestis.
 FH Key Location/Qualifiers
 FT CDS 17..532
 FT /*tag= a
 FT /*product= "F1 antigen"
 PN W09824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-333331/29.
 DR N-PSDB; V41596.
 PT Protection of animals against plague - using nucleic acid encoding

PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Disclosure: Pages 53-54; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 170 AA;

Query Match 32.6%; Score 170; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1e-153;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAALFGTATANAADLTASTATATLVEPARITLYEGAPITMDNGNIDT 83
 DB 1 MKRISVIAALFGTATANAADLTASTATATLVEPARITLYEGAPITMDNGNIDT 60
 QY 84 ELLVGTLLIGYKGTSTSVNFTDAAGDPMYLTFTSODGNHQTFTYIGKDSRDFDIS 143
 DB 61 ELLVGTLLIGYKGTSTSVNFTDAAGDPMYLTFTSODGNHQTFTYIGKDSRDFDIS 120
 QY 144 PVRNGENLVGDVYVATGSDOFFVRSIGSKGKLAGKTTDAVTYVSNQ 193
 DB 121 PVRNGENLVGDVYVATGSDOFFVRSIGSKGKLAGKTTDAVTYVSNQ 170

RESULT 9
 W01044

ID W01044 standard; Protein: 501 AA.

AC W01044:
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1/V antigen fusion.
 KW Plaque: vaccine; genetic immunisation; V antigen; lcrV;
 KW F1 antigen; cafi.
 OS Chimeric Yersinia pestis strain GB;
 OS Chimeric synthetic.
 FH Key location/Qualifiers
 FT 1. 21
 FT /label- sig_peptide
 FT /note- "F1 antigen signal peptide"
 FT 22. 170
 FT /label- F1 antigen
 FT /note- "mature F1 antigen"
 FT 171. 176
 FT /label- linker
 FT /note- "6-amino acid peptide linker"
 FT 176. 501
 FT /label- V antigen
 FT /note- "mature V antigen"
 FT protein

PN W09628551-A1.
 PD 19-SEP-1996
 PF 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED,
 DR WPTI 96-433824/43.
 DR N-PDB: T38249.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Example 3: Page 65-69; 98pp; English.
 CC A fusion protein (W01044) comprises the F1 antigen (see also
 CC W01043) and V antigen (see also W01041) of Yersinia pestis joined
 CC by a linker that allows each protein to attain its conformational
 CC state. It is the product of a gene fusion (T38249) obtd. by PCR
 CC amplification of Y. pestis DNA. F1/V fusion protein can be
 CC expressed by gut-colonising organism transformants, to induce
 CC an immune response against Y. pestis, the causative organism of
 CC plague.
 SQ Sequence 501 AA;

Query Match 32.6%; Score 170; DB 1; Length 501;
 Best Local Similarity 100.0%; Pred. No. 2.6e-153;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAALFGTATANAADLTASTATATLVEPARITLYEGAPITMDNGNIDT 83
 DB 1 MKRISVIAALFGTATANAADLTASTATATLVEPARITLYEGAPITMDNGNIDT 60
 QY 84 ELLVGTLLIGYKGTSTSVNFTDAAGDPMYLTFTSODGNHQTFTYIGKDSRDFDIS 143
 DB 61 ELLVGTLLIGYKGTSTSVNFTDAAGDPMYLTFTSODGNHQTFTYIGKDSRDFDIS 120
 QY 144 PVRNGENLVGDVYVATGSDOFFVRSIGSKGKLAGKTTDAVTYVSNQ 193
 DB 121 PVRNGENLVGDVYVATGSDOFFVRSIGSKGKLAGKTTDAVTYVSNQ 170

RESULT 10
 R79962

ID R79962 standard; Protein: 329 AA.

AC R79962:
 DT 18-APR-1996 (first entry)
 DE Partial lcrV (V antigen) of Y. pestis.
 KW lcrV; V antigen; virulence; plaque; vaccine; epitope.
 OS Yersinia pestis.
 PN W09524475-A1.
 PD 14-SEP-1995.
 PF 06-MAR-1995; G00481.
 PR 08-MAR-1994; GB-004577.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Leary SEC, Titball RW, Williamson ED, Leary SE;
 DR WPTI 95-328268/42.
 DR N-PDB: T04223.
 PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
 PT oral or parenteral vaccines for protection against plague
 PS Claim 6; Page 15-16; 25pp; English.
 CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrV) encoding
 CC all or a protective epitopic part of the mature V protein of Yersinia
 CC pestis. The protein was expressed as a fusion protein with maltose
 CC binding protein or glutathione-S-transferase in 3 different plasmid
 CC vectors. Y. pestis is the highly virulent causative organism of plague
 CC in a wide range of animals, including man. The V antigen (lcrV) is an
 CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
 CC The V antigen is postulated to act as a virulence antigen, and
 CC transformed microorganisms contg. recombinant DNA encoding a V antigen
 CC protein/peptide are useful in vaccines to protect against plague.
 SQ Sequence 329 AA;

Query Match 32.1%; Score 167; DB 1; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.4e-152;
 Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 LPAELKIVSVIOAEINKHLSSGTTINIDKSLNMDKNLYGTDEIRKASAEYLKRM 414
 DB 163 LPAELKIVSVIOAEINKHLSSGTTINIDKSLNMDKNLYGTDEIRKASAEYLKRM 222
 QY 415 POTTIOVDSEKRIYSIKDFLSEKRRKRGALGNLANSYSYNKNNELSHFATTCSDKSRP 474
 DB 223 POTTIOVDSEKRIYSIKDFLSEKRRKRGALGNLANSYSYNKNNELSHFATTCSDKSRP 282
 QY 475 LMDVYSOKTQSDTTSRNSAIEALNRFQKYDSVMORLDDTIGSK 521
 DB 283 LMDVYSOKTQSDTTSRNSAIEALNRFQKYDSVMORLDDTIGSK 329

RESULT 11
 W59788

ID W59788 standard; Protein: 149 AA.

AC W59788:
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen pYpF1mat149.

KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996: US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41609.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10: Page 63; 75pp: English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 149 AA;

Query Match 28.6%; Score 149; DB 1; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.4e-135;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 ADLSTATATATVEPARITLYKKGAPITMDNGNIDTELLVGLTLGGYKGTSTSV 104
 ID W59787
 AC W59787 standard; Protein; 150 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pyfmat150.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996: US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41609.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10: Page 63; 75pp: English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 149 AA;

RESULT 12

ID W59787
 AC W59787 standard; Protein; 150 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pyfmat150.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996: US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41609.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10: Page 63; 75pp: English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 150 AA;

Query Match 28.6%; Score 149; DB 1; Length 150;
 Best Local Similarity 100.0%; Pred. No. 1.4e-135;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 ADLSTATATATVEPARITLYKKGAPITMDNGNIDTELLVGLTLGGYKGTSTSV 104
 ID W59787
 AC W59787 standard; Protein; 150 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pyfmat150.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996: US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41609.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10: Page 63; 75pp: English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 150 AA;

QY 105 NPTDAADPMYLTFTSODGNHOFITKIGKDSRDPDISPKVNGENLVGDVYLATGSOD 164
 ID W59785
 AC W59785 standard; Protein; 192 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pyfmat192.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996: US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41598.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10: Pages 56-57; 75pp: English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 192 AA;

RESULT 13

ID W59785
 AC W59785 standard; Protein; 192 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pyfmat192.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996: US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41598.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10: Pages 56-57; 75pp: English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 192 AA;

Query Match 25.7%; Score 134; DB 1; Length 192;
 Best Local Similarity 100.0%; Pred. No. 4.5e-121;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKRISVIAIALFETITANAAADLTASTATATVEPARITLYKKGAPITMDNGNIDT 83
 ID W59785
 AC W59785 standard; Protein; 192 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pyfmat192.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996: US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41598.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10: Pages 56-57; 75pp: English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 192 AA;

RESULT 14

ID W59785
 AC W59785 standard; Protein; 192 AA.
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pyfmat192.
 KM F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997: U22617.
 PR 04-DEC-1996: US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI: 98-33331/29.
 N-PSDB: V41598.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10: Pages 56-57; 75pp: English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 192 AA;

PS Disclosure: Page 16; 27pp; English.
 CC The sequence represents the Y. pestis cafi (F1) antigen expressed
 CC from plasmid pGAL2a. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated Salmonella typhimurium or Salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 151 AA;

Query Match

24.6%; Score 128; DB 1; Length 151;

Best Local Similarity 100.0%; Pred. No. 2.1e-115; Mismatches 0; Indels 0; Gaps 0;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 TYKGAFTTMDNGNIDTELVGTLTGKTKGTSTSVNFTDAAGDPMTLFTSODGNN 125
 DB 24 TYKGAFTTMDNGNIDTELVGTLTGKTKGTSTSVNFTDAAGDPMTLFTSODGNN 83
 QY 126 HOFTTKVIGKDSRDFDISPRVNGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKRTDA 185
 DB 84 HOFTTKVIGKDSRDFDISPRVNGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKRTDA 143
 QY 186 VTVVSNQ 193
 DB 144 VTVVSNQ 151

RESULT 15

R76527
 ID R76527 standard; Protein; 151 AA.
 AC R76527:
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis cafi (F1) antigen.
 KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 KM bubonic plague; pneumonic plague.
 OS Yersinia pestis.
 PN MO9518231-A1.
 PD 06-JUL-1995.
 PE 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
 DR WPI: 95-246396/32.
 DR N-PSDB; Q92818.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure: Page 18; 27pp; English.
 CC The sequence represents the Y. pestis cafi (F1) antigen expressed
 CC from plasmid pGAL2a. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated Salmonella typhimurium or Salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 151 AA;

Query Match

24.6%; Score 128; DB 1; Length 151;

Best Local Similarity 100.0%; Pred. No. 2.1e-115; Mismatches 0; Indels 0; Gaps 0;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 TYKGAFTTMDNGNIDTELVGTLTGKTKGTSTSVNFTDAAGDPMTLFTSODGNN 125
 DB 24 TYKGAFTTMDNGNIDTELVGTLTGKTKGTSTSVNFTDAAGDPMTLFTSODGNN 83
 QY 126 HOFTTKVIGKDSRDFDISPRVNGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKRTDA 185

DB 84 HOFTTKVIGKDSRDFDISPRVNGENLVGDDVVLATGSODFFVRSIGSKGKLAAGKRTDA 143

QY 186 VTVVSNQ 193

DB 144 VTVVSNQ 151

Search completed: August 22, 2000, 17:41:58
 Job time: 2378 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:23:40 ; Search time 22.14 Seconds

(Without alignments)
360.727 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGNHHHHHHSSGHIDDD.....RFLQKXDSVMQRLLDPTSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 152396 seqs, 15239161 residues

Word size : 0

Total number of hits satisfying chosen parameters: 152396

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2-6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2-6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2-6/ptodata/1/1aa/6.COMB.pep:*
- 4: /cgn2-6/ptodata/1/1aa/PCrUS.COMB.pep:*
- 5: /cgn2-6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243	46.6	329	2	US-08-913-477-2
2	243	46.6	329	2	US-08-913-477-4
3	243	46.6	480	2	US-08-913-477-17
4	243	46.6	501	2	US-08-913-477-23
5	131	25.1	170	2	US-08-913-477-21
6	128	24.6	151	2	US-08-913-477-11
7	24	4.6	24	3	US-08-780-496-8
8	17	3.3	21	2	US-08-651-818A-21
9	17	3.3	23	1	US-08-480-604A-24
10	17	3.3	23	2	US-08-405-496A-24
11	17	3.3	53	2	US-08-651-818A-19
12	17	3.3	54	2	US-08-651-818A-23
13	17	3.3	323	1	US-08-591-989-5
14	17	3.3	462	1	US-08-480-604A-26
15	17	3.3	462	2	US-08-405-496A-26
16	16	3.1	254	3	US-09-167-434-7
17	12	2.3	17	2	US-08-600-783-13
18	11	2.1	302	4	US-08-203-532F-4
19	11	2.1	302	4	PCT-US95-01882A-4
20	11	2.1	303	2	US-08-203-532F-2
21	11	2.1	303	2	PCT-US95-01882A-2
22	11	2.1	414	4	PCT-US92-06840-2
23	10	1.9	449	2	US-08-927-394-2
24	10	1.9	480	1	US-07-883-292-2
25	10	1.9	480	2	US-08-331-644-2
26	10	1.9	480	2	PCT-US93-04102-2
27	10	1.9	763	2	US-08-677-862-2
28	10	1.9	763	2	US-09-252-571-2

29	10	1.9	2265	2	US-08-149-097D-36	Sequence 36, Appl
30	10	1.9	2509	2	US-08-149-097D-35	Sequence 35, Appl
31	9	1.7	17	3	US-08-946-475-10	Sequence 10, Appl
32	9	1.7	20	1	US-08-373-134D-3	Sequence 3, Appl
33	9	1.7	20	2	US-09-114-637-3	Sequence 3, Appl
34	9	1.7	20	2	US-08-485-942A-99	Sequence 99, Appl
35	9	1.7	21	2	US-08-432-871C-46	Sequence 46, Appl
36	9	1.7	21	2	US-08-712-878-6	Sequence 6, Appl
37	9	1.7	21	3	US-08-485-942A-98	Sequence 98, Appl
38	9	1.7	43	2	US-08-347-563A-12	Sequence 12, Appl
39	9	1.7	43	2	US-08-485-942A-12	Sequence 12, Appl
40	9	1.7	110	1	US-08-434-705B-15	Sequence 15, Appl
41	9	1.7	110	1	US-09-086-201-15	Sequence 15, Appl
42	9	1.7	408	2	US-08-683-007A-2	Sequence 2, Appl
43	9	1.7	420	2	US-08-845-998-8	Sequence 8, Appl
44	9	1.7	711	3	US-08-946-475-9	Sequence 9, Appl
45	8	1.5	10	1	US-07-807-529A-73	Sequence 7, Appl
46	8	1.5	10	2	US-08-482-142-150	Sequence 150, App
47	8	1.5	10	2	US-08-478-572-150	Sequence 150, App
48	8	1.5	10	3	US-08-300-928C-88	Sequence 88, Appl
49	8	1.5	10	3	US-08-430-944D-88	Sequence 88, Appl
50	8	1.5	10	3	US-08-430-014-68	Sequence 88, Appl
51	8	1.5	10	4	PCT-US94-14073-17	Sequence 17, Appl
52	8	1.5	14	1	US-07-807-529A-76	Sequence 76, Appl
53	8	1.5	14	3	US-08-300-928C-91	Sequence 91, Appl
54	8	1.5	14	3	US-08-430-944D-91	Sequence 91, Appl
55	8	1.5	14	3	US-08-430-014-91	Sequence 91, Appl
56	8	1.5	15	2	US-08-467-603-53	Sequence 53, Appl
57	8	1.5	15	2	US-08-466-793-53	Sequence 53, Appl
58	8	1.5	15	2	US-08-491-861A-53	Sequence 53, Appl
59	8	1.5	19	1	US-08-346-293-8	Sequence 8, Appl
60	8	1.5	23	1	US-08-346-293-8	Sequence 8, Appl
61	8	1.5	27	1	US-08-346-293-8	Sequence 8, Appl
62	8	1.5	31	1	US-08-346-293-23	Sequence 23, Appl
63	8	1.5	38	1	US-08-346-293-23	Sequence 23, Appl
64	8	1.5	42	1	US-08-346-293-25	Sequence 25, Appl
65	8	1.5	60	1	US-08-255-457-1	Sequence 1, Appl
66	8	1.5	60	2	US-09-115-032-1	Sequence 1, Appl
67	8	1.5	60	4	PCT-US95-05772-1	Sequence 1, Appl
68	8	1.5	96	1	US-07-807-529A-39	Sequence 39, Appl
69	8	1.5	96	3	US-08-430-944D-103	Sequence 103, App
70	8	1.5	221	1	US-08-621-081A-19	Sequence 19, Appl
71	8	1.5	494	1	US-08-485-618-103	Sequence 103, App
72	8	1.5	494	2	US-08-605-672-103	Sequence 103, App
73	8	1.5	494	2	US-08-482-293A-103	Sequence 103, App
74	8	1.5	494	2	US-08-943-363-103	Sequence 103, App
75	7	1.3	7	1	US-08-100-744-7	Sequence 7, Appl
76	7	1.3	7	1	US-08-284-784-7	Sequence 7, Appl
77	7	1.3	7	2	US-08-854-811-7	Sequence 7, Appl
78	7	1.3	8	1	US-08-346-293-5	Sequence 5, Appl
79	7	1.3	9	1	US-08-155-171B-4	Sequence 4, Appl
80	7	1.3	9	2	US-08-435-998-4	Sequence 4, Appl
81	7	1.3	12	1	US-08-346-293-6	Sequence 6, Appl
82	7	1.3	14	2	US-08-448-418-100	Sequence 100, App
83	7	1.3	17	1	US-08-021-603A-19	Sequence 19, Appl
84	7	1.3	17	1	US-08-155-171B-37	Sequence 37, Appl
85	7	1.3	17	4	US-08-435-998-37	Sequence 37, Appl
86	7	1.3	17	4	PCT-US94-01176-19	Sequence 19, Appl
87	7	1.3	45	1	US-08-056-200-97	Sequence 97, Appl
88	7	1.3	45	2	US-08-800-644-57	Sequence 57, Appl
89	7	1.3	88	1	US-08-690-011A-11	Sequence 11, Appl
90	7	1.3	91	2	US-07-987-272A-11	Sequence 11, Appl
91	7	1.3	92	2	US-08-918-727-5	Sequence 5, Appl
92	7	1.3	92	2	US-09-051-589-1	Sequence 1, Appl
93	7	1.3	98	2	US-08-690-011A-19	Sequence 19, Appl
94	7	1.3	105	2	US-08-690-011A-15	Sequence 15, Appl
95	7	1.3	115	2	US-08-690-011A-21	Sequence 21, Appl
96	7	1.3	124	1	US-08-434-705B-16	Sequence 16, Appl
97	7	1.3	124	1	US-08-086-201-16	Sequence 16, Appl
98	7	1.3	210	1	US-08-234-783-2	Sequence 2, Appl
99	7	1.3	210	4	US-08-456-907-2	Sequence 2, Appl
100	7	1.3	210	4	PCT-US95-05523-2	Sequence 2, Appl
101	7	1.3	249	1	US-08-155-171B-32	Sequence 32, Appl

102	7	1	3	249	2	US-08-435-998-32	Sequence 32, Appl	175	6	1	2	12	4	PCT-US93-08062-48	Sequence 48, Appl
103	7	1	3	302	2	US-08-893-853-3	Sequence 3, Appl	176	6	1	2	12	4	PCT-US94-09700-4	Sequence 4, Appl
104	7	1	3	368	1	US-08-211-942-17	Sequence 17, Appl	177	6	1	2	13	1	US-08-571-643A-4	Sequence 22, Appl
105	7	1	3	331	1	US-08-589-981-2	Sequence 2, Appl	178	6	1	2	13	1	US-08-439-747A-2	Sequence 22, Appl
106	7	1	3	685	3	US-08-878-989-1	Sequence 2, Appl	179	6	1	2	13	1	US-08-482-577B-15	Sequence 15, Appl
107	7	1	3	685	3	US-09-136-282-2	Sequence 2, Appl	180	6	1	2	13	2	US-08-440-409B-32	Sequence 22, Appl
108	7	1	3	946	3	US-08-560-005-4	Sequence 4, Appl	181	6	1	2	14	1	US-08-226-264-7	Sequence 7, Appl
109	7	1	3	1489	5	5183745-2	Patent No. 5183745	182	6	1	2	14	1	US-08-226-264-19	Sequence 19, Appl
110	7	1	3	1794	5	5183745-6	Patent No. 5183745	183	6	1	2	14	1	US-08-226-264-20	Sequence 20, Appl
111	7	1	3	3729	5	US-08-804-227C-4	Sequence 6, Appl	184	6	1	2	14	2	US-08-467-603-58	Sequence 58, Appl
112	6	1	2	6	1	US-08-160-670A-6	Sequence 6, Appl	185	6	1	2	14	2	US-08-466-793-58	Sequence 58, Appl
113	6	1	2	6	1	US-08-140-222-1	Sequence 1, Appl	186	6	1	2	14	2	US-08-491-861A-58	Sequence 58, Appl
114	6	1	2	6	1	US-08-131-365B-48	Sequence 48, Appl	187	6	1	2	14	2	US-08-672-345C-40	Sequence 40, Appl
115	6	1	2	6	1	US-08-469-486-47	Sequence 47, Appl	188	6	1	2	15	1	US-08-469-486-48	Sequence 48, Appl
116	6	1	2	6	1	US-08-460-343B-51	Sequence 51, Appl	189	6	1	2	15	2	US-08-469-658-48	Sequence 48, Appl
117	6	1	2	6	1	US-08-289-567-1	Sequence 1, Appl	190	6	1	2	16	1	US-08-218-025A-111	Sequence 11, App
118	6	1	2	6	1	US-08-398-028B-51	Sequence 51, Appl	191	6	1	2	16	1	US-08-578-649-13	Sequence 13, Appl
119	6	1	2	6	1	US-08-660-626-4	Sequence 4, Appl	192	6	1	2	16	1	US-08-226-264-8	Sequence 8, Appl
120	6	1	2	6	1	US-08-434-705B-17	Sequence 17, Appl	193	6	1	2	16	1	US-08-226-264-21	Sequence 21, Appl
121	6	1	2	6	1	US-08-470-837-22	Sequence 22, Appl	194	6	1	2	16	4	PCT-US94-04361-56	Sequence 56, Appl
122	6	1	2	6	2	US-08-472-244-14	Sequence 14, Appl	195	6	1	2	17	1	US-08-323-445A-18	Sequence 18, Appl
123	6	1	2	6	2	US-08-504-265B-51	Sequence 51, Appl	196	6	1	2	17	1	US-08-515-903A-18	Sequence 18, Appl
124	6	1	2	6	2	US-08-623-833B-16	Sequence 16, Appl	197	6	1	2	17	3	US-09-136-421-10	Sequence 10, Appl
125	6	1	2	6	2	US-08-652-507-8	Sequence 8, Appl	198	6	1	2	17	4	PCT-US95-12840-18	Sequence 18, Appl
126	6	1	2	6	2	US-08-668-123-48	Sequence 48, Appl	199	6	1	2	17	5	3428012-6	Patent No. 3428012
127	6	1	2	6	2	US-08-469-658-47	Sequence 47, Appl	200	6	1	2	17	5	3451506-6	Patent No. 3451506
128	6	1	2	6	2	US-09-086-201-17	Sequence 17, Appl	201	6	1	2	18	1	US-08-145-006C-6	Sequence 6, Appl
129	6	1	2	6	2	US-08-591-196-57	Sequence 57, Appl	202	6	1	2	18	1	US-08-578-649-14	Sequence 14, Appl
130	6	1	2	6	3	US-08-996-139-18	Sequence 18, Appl	203	6	1	2	18	2	US-08-356-060A-47	Sequence 47, Appl
131	6	1	2	6	3	US-08-481-435-42	Sequence 42, Appl	204	6	1	2	18	2	US-08-670-175-8	Sequence 8, Appl
132	6	1	2	6	3	US-08-772-440-28	Sequence 28, Appl	205	6	1	2	18	3	US-08-895-707-15	Sequence 15, Appl
133	6	1	2	6	4	PCT-US94-04361-55	Sequence 55, Appl	206	6	1	2	19	4	PCT-US94-00545-6	Sequence 6, Appl
134	6	1	2	6	4	US-08-762-433-6	Sequence 6, Appl	207	6	1	2	19	4	US-08-462-436-13	Sequence 13, Appl
135	6	1	2	6	4	US-09-001-219-6	Sequence 6, Appl	208	6	1	2	19	3	US-09-167-434-15	Sequence 15, Appl
136	6	1	2	6	4	US-08-481-658B-22	Sequence 22, Appl	209	6	1	2	20	1	US-08-416-336-4	Sequence 4, Appl
137	6	1	2	6	4	US-08-477-504A-22	Sequence 22, Appl	210	6	1	2	20	2	US-08-543-020-12	Sequence 12, Appl
138	6	1	2	6	4	US-08-486-755A-22	Sequence 22, Appl	211	6	1	2	21	2	US-08-480-473B-16	Sequence 16, Appl
139	6	1	2	6	4	US-08-485-862B-22	Sequence 22, Appl	212	6	1	2	21	3	US-08-915-213-46	Sequence 46, Appl
140	6	1	2	6	4	US-08-787-739-22	Sequence 22, Appl	213	6	1	2	22	1	US-08-127-909-19	Sequence 19, Appl
141	6	1	2	6	4	US-08-487-077A-22	Sequence 22, Appl	214	6	1	2	22	2	US-08-457-353-14	Sequence 14, Appl
142	6	1	2	6	4	US-08-816-977-11	Sequence 11, Appl	215	6	1	2	23	2	US-08-273-146-37	Sequence 37, Appl
143	6	1	2	6	4	US-08-100-744-12	Sequence 12, Appl	216	6	1	2	23	2	US-08-350-260A-42	Sequence 42, Appl
144	6	1	2	6	4	US-08-284-784-12	Sequence 12, Appl	217	6	1	2	24	1	US-08-021-603A-16	Sequence 16, Appl
145	6	1	2	6	4	US-08-323-445A-2	Sequence 2, Appl	218	6	1	2	24	4	PCT-US94-01176-16	Sequence 16, Appl
146	6	1	2	6	4	US-08-515-503A-2	Sequence 2, Appl	219	6	1	2	25	4	US-08-127-909-19	Sequence 19, Appl
147	6	1	2	6	4	US-08-737-757-12	Sequence 12, Appl	220	6	1	2	25	2	US-08-457-353-19	Sequence 19, Appl
148	6	1	2	6	4	US-08-854-811-12	Sequence 12, Appl	221	6	1	2	25	2	US-08-473-475A-21	Sequence 21, Appl
149	6	1	2	6	4	US-09-143-470-19	Sequence 19, Appl	222	6	1	2	26	1	US-08-227-372-2	Sequence 2, Appl
150	6	1	2	6	4	US-09-254-567-1	Sequence 1, Appl	223	6	1	2	26	2	US-08-657-641-12	Sequence 12, Appl
151	6	1	2	6	4	US-08-885-707-9	Sequence 9, Appl	224	6	1	2	26	3	US-08-470-397-2	Sequence 2, Appl
152	6	1	2	6	4	PCT-US95-12840-2	Sequence 2, Appl	225	6	1	2	26	3	US-08-895-707-21	Sequence 21, Appl
153	6	1	2	6	4	US-08-340-812-16	Sequence 16, Appl	226	6	1	2	26	4	PCT-US94-07233-12	Sequence 12, Appl
154	6	1	2	6	4	US-08-459-064B-16	Sequence 16, Appl	227	6	1	2	29	4	PCT-US94-05150-37	Sequence 37, Appl
155	6	1	2	6	4	US-08-737-757-10	Sequence 10, Appl	228	6	1	2	31	2	US-08-663-227-37	Sequence 37, Appl
156	6	1	2	6	4	US-08-628-665-14	Sequence 14, Appl	229	6	1	2	32	1	US-08-323-445A-17	Sequence 17, Appl
157	6	1	2	6	4	US-08-524-757-30	Sequence 30, Appl	230	6	1	2	32	1	US-08-515-903A-17	Sequence 17, Appl
158	6	1	2	6	4	US-08-524-757-30	Sequence 30, Appl	231	6	1	2	32	3	US-08-737-336-1	Sequence 1, Appl
159	6	1	2	6	4	US-08-460-421A-16	Sequence 16, Appl	232	6	1	2	32	4	PCT-US95-12840-17	Sequence 17, Appl
160	6	1	2	6	4	US-08-717-169-13	Sequence 13, Appl	233	6	1	2	36	3	US-08-554-385-23	Sequence 23, Appl
161	6	1	2	6	4	US-08-220-40A-3	Sequence 3, Appl	234	6	1	2	36	3	US-08-679-865-39	Sequence 39, Appl
162	6	1	2	6	4	US-08-792-553-24	Sequence 24, Appl	235	6	1	2	39	2	US-08-680-876-39	Sequence 39, Appl
163	6	1	2	6	4	PCT-US93-00909-13	Sequence 13, Appl	236	6	1	2	40	2	US-08-273-146-39	Sequence 39, Appl
164	6	1	2	6	4	PCT-US95-04433-9	Sequence 9, Appl	237	6	1	2	42	1	US-08-453-924-8	Sequence 8, Appl
165	6	1	2	6	4	US-08-133-804-11	Sequence 11, Appl	238	6	1	2	42	1	US-08-664-449-39	Sequence 39, Appl
166	6	1	2	6	4	US-08-461-838-11	Sequence 11, Appl	239	6	1	2	43	2	US-08-273-146-41	Sequence 41, Appl
167	6	1	2	6	4	US-08-737-757-4	Sequence 4, Appl	240	6	1	2	43	2	US-08-652-816A-43	Sequence 43, Appl
168	6	1	2	6	4	US-08-461-386-11	Sequence 11, Appl	241	6	1	2	44	2	US-08-652-367-6	Sequence 6, Appl
169	6	1	2	6	4	US-08-100-744-11	Sequence 11, Appl	242	6	1	2	45	1	US-08-220-378-1	Sequence 1, Appl
170	6	1	2	6	4	US-08-284-784-11	Sequence 11, Appl	243	6	1	2	45	2	US-08-696-012-1	Sequence 1, Appl
171	6	1	2	6	4	US-08-459-064B-35	Sequence 35, Appl	244	6	1	2	45	2	US-08-889-291-3	Sequence 3, Appl
172	6	1	2	6	4	US-08-460-421A-35	Sequence 35, Appl	245	6	1	2	48	1	US-08-457-245-21	Sequence 21, Appl
173	6	1	2	6	4	US-08-854-811-11	Sequence 11, Appl	246	6	1	2	55	1	US-08-466-033-273	Sequence 273, App
174	6	1	2	6	4	US-08-429-964-48	Sequence 48, Appl	247	6	1	2	55	2	US-08-444-733-273	Sequence 273, App

248	6	1.2	55	2	US-08-464-134-273	Sequence 273, App	321	6	1.2	233	3	US-08-837-056-9	Sequence 9, Appl
249	6	1.2	55	2	US-08-461-361-273	Sequence 273, App	322	6	1.2	242	3	US-09-167-434-8	Sequence 8, Appl
250	6	1.2	55	2	US-08-485-910-273	Sequence 273, App	323	6	1.2	243	3	US-09-167-434-8	Sequence 9, Appl
251	6	1.2	57	2	US-08-652-816A-41	Sequence 41, Appl	324	6	1.2	244	3	US-08-480-753-4	Sequence 4, Appl
252	6	1.2	57	2	US-08-958-201-18	Sequence 18, Appl	325	6	1.2	246	3	US-09-167-434-11	Sequence 11, Appl
253	6	1.2	67	2	US-08-142-551B-9	Sequence 9, Appl	326	6	1.2	248	4	PCT-US93-03895-6	Sequence 6, Appl
254	6	1.2	69	2	US-08-687-865A-21	Sequence 21, Appl	327	6	1.2	250	4	PCT-US93-03895-6	Sequence 6, Appl
255	6	1.2	71	3	US-08-783-974-6	Sequence 6, Appl	328	6	1.2	251	1	US-08-397-613A-74	Sequence 14, Appl
256	6	1.2	79	3	US-08-448-418-97	Sequence 97, Appl	329	6	1.2	252	1	US-08-792-824-10	Sequence 13, Appl
257	6	1.2	81	2	US-08-717-169-15	Sequence 15, Appl	330	6	1.2	252	3	US-08-988-813-13	Sequence 13, Appl
258	6	1.2	85	3	US-08-772-440-32	Sequence 32, Appl	331	6	1.2	254	2	US-08-792-824-7	Sequence 4, Appl
259	6	1.2	87	1	US-08-160-670A-7	Sequence 7, Appl	332	6	1.2	254	2	US-08-792-824-7	Sequence 4, Appl
260	6	1.2	89	1	US-08-340-812-13	Sequence 13, Appl	333	6	1.2	254	2	US-08-792-824-10	Sequence 10, Appl
261	6	1.2	89	1	US-08-459-064B-13	Sequence 13, Appl	334	6	1.2	254	2	US-08-792-824-13	Sequence 13, Appl
262	6	1.2	89	2	US-08-460-421A-13	Sequence 13, Appl	335	6	1.2	255	3	US-09-167-434-10	Sequence 10, Appl
263	6	1.2	89	2	US-08-717-169-7	Sequence 7, Appl	336	6	1.2	257	3	US-08-486-099-113	Sequence 113, App
264	6	1.2	89	3	US-08-946-026-2	Sequence 2, Appl	337	6	1.2	257	3	US-08-360-107A-123	Sequence 123, App
265	6	1.2	93	2	US-08-690-011A-9	Sequence 9, Appl	338	6	1.2	257	3	US-08-484-223B-113	Sequence 113, App
266	6	1.2	94	2	US-08-266-451B-20	Sequence 20, Appl	339	6	1.2	257	3	US-08-919-597-113	Sequence 113, App
267	6	1.2	94	2	US-08-748-725-20	Sequence 20, Appl	340	6	1.2	257	3	US-08-475-668A-113	Sequence 113, App
268	6	1.2	97	3	US-08-816-977-23	Sequence 23, Appl	341	6	1.2	257	3	US-08-485-551A-113	Sequence 113, App
269	6	1.2	97	3	US-08-816-977-27	Sequence 27, Appl	342	6	1.2	261	1	US-08-160-670A-45	Sequence 45, App
270	6	1.2	98	3	US-08-478-097A-4	Sequence 4, Appl	343	6	1.2	262	1	US-08-397-633A-73	Sequence 73, Appl
271	6	1.2	99	2	US-08-672-345C-94	Sequence 94, Appl	344	6	1.2	263	1	US-08-160-670A-32	Sequence 32, Appl
272	6	1.2	99	2	US-08-484-905-62	Sequence 62, Appl	345	6	1.2	263	2	US-08-353-476-66	Sequence 66, Appl
273	6	1.2	99	3	US-08-481-985B-62	Sequence 62, Appl	346	6	1.2	263	2	US-08-353-476-67	Sequence 67, Appl
274	6	1.2	109	2	US-08-672-345C-4	Sequence 4, Appl	347	6	1.2	264	2	US-08-436-748-8	Sequence 8, Appl
275	6	1.2	112	1	US-08-211-202-133	Sequence 133, App	348	6	1.2	264	2	US-07-857-224B-9	Sequence 9, Appl
276	6	1.2	116	1	US-08-478-039-67	Sequence 67, Appl	349	6	1.2	265	3	US-08-483-857-8	Sequence 8, Appl
277	6	1.2	116	1	US-08-476-349A-67	Sequence 67, Appl	350	6	1.2	275	2	US-08-578-709-13	Sequence 13, Appl
278	6	1.2	116	1	US-08-481-377-26	Sequence 26, Appl	351	6	1.2	281	1	US-08-397-633A-75	Sequence 75, Appl
279	6	1.2	116	2	US-08-491-835-24	Sequence 24, Appl	352	6	1.2	287	1	US-08-397-633A-76	Sequence 76, Appl
280	6	1.2	116	3	US-09-153-733A-26	Sequence 26, Appl	353	6	1.2	297	1	US-08-011-398B-4	Sequence 4, Appl
281	6	1.2	116	4	PCT-US94-00666-26	Sequence 26, Appl	354	6	1.2	297	1	US-08-464-051-4	Sequence 4, Appl
282	6	1.2	116	4	PCT-US94-00666-24	Sequence 24, Appl	355	6	1.2	297	2	US-08-462-498-4	Sequence 4, Appl
283	6	1.2	120	4	PCT-US93-03895-2	Sequence 2, Appl	356	6	1.2	297	2	US-08-554-385-3	Sequence 3, Appl
284	6	1.2	121	4	PCT-US93-03895-12	Sequence 12, Appl	357	6	1.2	298	3	US-09-080-004-9	Sequence 9, Appl
285	6	1.2	126	1	US-08-268-348A-14	Sequence 14, Appl	358	6	1.2	301	2	US-08-661-051-14	Sequence 14, Appl
286	6	1.2	127	1	US-08-253-155A-35	Sequence 35, Appl	359	6	1.2	304	1	US-08-132-405-4	Sequence 4, Appl
287	6	1.2	128	1	US-08-426-627-14	Sequence 14, Appl	360	6	1.2	304	1	US-08-393-938A-4	Sequence 4, Appl
288	6	1.2	133	1	US-08-268-348A-8	Sequence 8, Appl	361	6	1.2	304	4	PCT-US91-01861-4	Sequence 4, Appl
289	6	1.2	133	1	US-08-668-348A-10	Sequence 10, Appl	362	6	1.2	308	2	US-08-499-958-2	Sequence 2, Appl
290	6	1.2	172	2	US-08-557-309B-39	Sequence 39, Appl	363	6	1.2	308	2	US-08-793-565-12	Sequence 12, Appl
291	6	1.2	172	3	US-08-834-306-39	Sequence 39, Appl	364	6	1.2	311	3	US-08-946-914-4	Sequence 4, Appl
292	6	1.2	174	3	US-08-906-136A-2	Sequence 2, Appl	365	6	1.2	312	1	US-08-414-625-2	Sequence 2, Appl
293	6	1.2	174	3	US-08-806-136A-4	Sequence 4, Appl	366	6	1.2	315	2	US-08-484-956-91	Sequence 91, Appl
294	6	1.2	179	4	US-08-809-267-7	Sequence 7, Appl	367	6	1.2	315	3	US-08-757-653-91	Sequence 91, Appl
295	6	1.2	179	4	PCT-US95-13662A-7	Sequence 7, Appl	368	6	1.2	319	3	US-08-630-172-12	Sequence 12, Appl
296	6	1.2	180	3	US-08-772-440-31	Sequence 31, Appl	369	6	1.2	320	2	US-08-757-653-163	Sequence 163, App
297	6	1.2	183	5	5168049-3	Patent No. 5168049	370	6	1.2	320	2	US-08-823-516-61	Sequence 61, Appl
298	6	1.2	184	3	US-09-023-082A-30	Sequence 30, Appl	371	6	1.2	323	3	US-08-816-977-21	Sequence 21, Appl
299	6	1.2	197	2	US-08-879-561-5	Sequence 5, Appl	372	6	1.2	326	2	US-08-816-977-25	Sequence 25, Appl
300	6	1.2	198	1	US-08-642-253-131	Sequence 131, App	373	6	1.2	326	2	US-08-816-977-25	Sequence 25, Appl
301	6	1.2	198	1	US-08-397-633A-41	Sequence 41, Appl	374	6	1.2	333	2	US-09-123-851-1	Sequence 1, Appl
302	6	1.2	198	1	US-08-397-633A-52	Sequence 52, Appl	375	6	1.2	343	2	US-08-447-430A-33	Sequence 33, Appl
303	6	1.2	214	2	US-08-846-790A-1	Sequence 1, Appl	376	6	1.2	344	1	US-07-941-523-24	Sequence 24, Appl
304	6	1.2	214	2	US-08-935-333-1	Sequence 1, Appl	377	6	1.2	351	2	US-08-701-191A-3	Sequence 3, Appl
305	6	1.2	218	1	US-08-644-664B-14	Sequence 14, Appl	378	6	1.2	357	2	US-08-674-612-3	Sequence 3, Appl
306	6	1.2	218	1	US-08-761-277-14	Sequence 14, Appl	379	6	1.2	360	1	US-08-469-421-14	Sequence 14, Appl
307	6	1.2	219	2	US-08-621-081A-16	Sequence 16, Appl	380	6	1.2	360	1	US-08-250-975-13	Sequence 13, Appl
308	6	1.2	223	1	US-08-430-633-1	Sequence 1, Appl	381	6	1.2	360	1	US-08-920-296-3	Sequence 3, Appl
309	6	1.2	223	1	US-08-336-854-1	Sequence 1, Appl	382	6	1.2	360	2	US-08-605-002A-14	Sequence 14, Appl
310	6	1.2	224	1	US-08-360-673-7	Sequence 7, Appl	383	6	1.2	360	2	US-08-950-448A-14	Sequence 14, Appl
311	6	1.2	224	1	US-08-414-625-4	Sequence 4, Appl	384	6	1.2	360	2	US-08-554-385-20	Sequence 20, Appl
312	6	1.2	227	5	5428012-3	Patent No. 5428012	385	6	1.2	360	4	PCT-US94-10599-14	Sequence 8, Appl
313	6	1.2	227	5	5428012-4	Patent No. 5428012	386	6	1.2	376	3	US-08-751-513-8	Sequence 3, Appl
314	6	1.2	227	5	5451506-3	Patent No. 5451506	387	6	1.2	377	2	US-08-929-922B-2	Sequence 2, Appl
315	6	1.2	228	5	5428012-2	Patent No. 5428012	388	6	1.2	378	2	US-08-707-237A-104	Sequence 104, App
316	6	1.2	228	5	5451506-2	Patent No. 5451506	389	6	1.2	378	2	US-08-642-246-26	Sequence 26, Appl
317	6	1.2	233	1	US-08-446-918A-4	Sequence 4, Appl	390	6	1.2	380	1	US-08-585-758A-4	Sequence 4, Appl
318	6	1.2	233	2	US-08-480-753-2	Sequence 2, Appl	391	6	1.2	380	1	US-08-977-818-4	Sequence 9, Appl
319	6	1.2	233	2	US-08-80-806-4	Sequence 4, Appl	392	6	1.2	380	1		
320	6	1.2	233	3	US-09-041-889-9	Sequence 9, Appl	393	6	1.2	380	1		

394	6	1.2	380	2	US-08-670-274B-4	Sequence 4, Appl	467	6	1.2	533	1	US-08-374-155A-14	Sequence 14, Appl
395	6	1.2	381	1	US-08-585-758A-2	Sequence 2, Appl	468	6	1.2	533	2	US-08-785-996-14	Sequence 14, Appl
396	6	1.2	381	1	US-08-977-818-2	Sequence 2, Appl	469	6	1.2	601	2	US-08-606-288-7	Sequence 10, Appl
397	6	1.2	381	2	US-08-670-274B-2	Sequence 2, Appl	470	6	1.2	601	2	US-08-606-288-10	Sequence 10, Appl
398	6	1.2	381	2	US-08-786-999-3	Sequence 3, Appl	471	6	1.2	612	2	US-08-752-907B-11	Sequence 11, Appl
399	6	1.2	386	3	US-08-895-707-7	Sequence 7, Appl	472	6	1.2	616	3	US-08-895-707-2	Sequence 2, Appl
400	6	1.2	390	2	US-08-786-999-1	Sequence 1, Appl	473	6	1.2	628	2	US-08-394-177-6	Sequence 6, Appl
401	6	1.2	391	1	US-07-995-657-6	Sequence 6, Appl	474	6	1.2	630	2	US-08-596-319-2	Sequence 2, Appl
402	6	1.2	391	1	US-08-474-587-6	Sequence 6, Appl	475	6	1.2	631	1	US-08-605-541B-12	Sequence 12, Appl
403	6	1.2	391	2	US-08-244-951A-10	Sequence 10, Appl	476	6	1.2	652	2	US-08-770-761A-5	Sequence 5, Appl
404	6	1.2	391	2	US-08-389-011-73	Sequence 23, Appl	477	6	1.2	657	3	US-08-480-640A-115	Sequence 115, App
405	6	1.2	391	3	US-08-403-917A-23	Sequence 23, Appl	478	6	1.2	657	3	US-08-480-640A-193	Sequence 193, App
406	6	1.2	391	4	PCT-US95-15696-2	Sequence 2, Appl	479	6	1.2	660	1	US-08-542-363-4	Sequence 4, Appl
407	6	1.2	393	4	US-08-977-554-2	Sequence 2, Appl	480	6	1.2	662	1	US-08-542-363-126	Sequence 126, App
408	6	1.2	401	5	5252556-1	Patent No. 5252556	481	6	1.2	662	1	US-08-397-633A-36	Sequence 36, Appl
409	6	1.2	403	2	US-09-061-337-10	Sequence 10, Appl	482	6	1.2	682	3	US-08-481-635-6	Sequence 6, Appl
410	6	1.2	403	2	US-09-123-139-10	Sequence 10, Appl	483	6	1.2	701	2	US-08-533-669A-16	Sequence 16, Appl
411	6	1.2	403	3	US-09-340-991-10	Sequence 10, Appl	484	6	1.2	705	2	US-08-770-761A-7	Sequence 7, Appl
412	6	1.2	406	2	US-08-222-719-1	Sequence 1, Appl	485	6	1.2	729	1	US-07-640-029-3	Sequence 3, Appl
413	6	1.2	406	2	US-08-470-925-1	Sequence 1, Appl	486	6	1.2	731	1	US-07-921-807B-5	Sequence 5, Appl
414	6	1.2	406	2	US-08-471-613-1	Sequence 1, Appl	487	6	1.2	731	1	US-08-441-944A-5	Sequence 5, Appl
415	6	1.2	406	2	PCT-US93-10443-1	Sequence 1, Appl	488	6	1.2	733	1	US-07-640-029-4	Sequence 4, Appl
416	6	1.2	407	2	US-08-776-585-3	Sequence 3, Appl	489	6	1.2	733	1	US-07-921-807B-6	Sequence 6, Appl
417	6	1.2	423	2	US-08-910-798-2	Sequence 2, Appl	490	6	1.2	733	1	US-08-441-944A-16	Sequence 16, Appl
418	6	1.2	423	2	US-08-313-533-13	Sequence 13, Appl	491	6	1.2	737	1	US-08-188-582-16	Sequence 16, Appl
419	6	1.2	425	3	US-08-767-993-13	Sequence 13, Appl	492	6	1.2	737	1	US-08-646-715-16	Sequence 16, Appl
420	6	1.2	427	2	US-08-846-021A-8	Sequence 8, Appl	493	6	1.2	761	2	US-08-849-480A-9	Sequence 9, Appl
421	6	1.2	439	1	US-08-507-431-4	Sequence 4, Appl	494	6	1.2	762	1	US-08-707-237A-84	Sequence 84, Appl
422	6	1.2	439	2	US-08-700-546-2	Sequence 2, Appl	495	6	1.2	762	1	US-08-642-255-114	Sequence 114, App
423	6	1.2	439	3	US-08-902-655A-4	Sequence 4, Appl	496	6	1.2	762	1	US-08-642-255-120	Sequence 120, App
424	6	1.2	439	3	US-09-116-622-4	Sequence 4, Appl	497	6	1.2	762	1	US-08-397-633A-26	Sequence 26, App
425	6	1.2	450	2	US-08-665-037-2	Sequence 4, Appl	498	6	1.2	762	1	US-08-397-633A-31	Sequence 31, Appl
426	6	1.2	450	2	US-08-666-067-2	Sequence 2, Appl	499	6	1.2	764	1	US-08-424-567-2	Sequence 2, Appl
427	6	1.2	450	2	US-08-733-870-2	Sequence 2, Appl	500	6	1.2	764	2	US-08-511-928-2	Sequence 2, Appl
428	6	1.2	458	5	5217891-15	Patent No. 5217891	501	6	1.2	768	2	US-08-560-398-2	Sequence 2, Appl
429	6	1.2	474	2	US-09-096-982-8	Sequence 8, Appl	502	6	1.2	770	1	US-08-369-796-12	Sequence 12, Appl
430	6	1.2	474	2	US-08-397-633A-8	Sequence 8, Appl	503	6	1.2	770	1	US-08-416-581B-1	Sequence 1, Appl
431	6	1.2	479	1	US-08-357-633A-78	Sequence 78, Appl	504	6	1.2	770	1	US-08-416-581B-5	Sequence 5, Appl
432	6	1.2	481	2	US-08-477-451-19	Sequence 19, Appl	505	6	1.2	770	1	US-08-416-581B-9	Sequence 9, Appl
433	6	1.2	486	1	US-08-397-633A-77	Sequence 77, Appl	506	6	1.2	770	2	US-08-852-091-12	Sequence 12, Appl
434	6	1.2	494	4	US-08-464-340A-4	Sequence 4, Appl	507	6	1.2	770	2	US-08-820-754-12	Sequence 12, Appl
435	6	1.2	494	4	PCT-US94-08449A-4	Sequence 4, Appl	508	6	1.2	770	3	US-08-956-652-12	Sequence 12, Appl
436	6	1.2	498	2	US-07-916-098A-4	Sequence 4, Appl	509	6	1.2	770	4	PCT-US95-17025-12	Sequence 12, Appl
437	6	1.2	500	2	US-08-578-709-15	Sequence 15, Appl	510	6	1.2	771	1	US-08-276-099A-14	Sequence 14, Appl
438	6	1.2	504	2	US-08-896-005-3	Sequence 3, Appl	511	6	1.2	771	1	US-08-781-890-14	Sequence 14, Appl
439	6	1.2	511	3	US-09-105-039A-2	Sequence 2, Appl	512	6	1.2	771	1	US-08-471-396A-4	Sequence 4, Appl
440	6	1.2	514	2	US-08-865-311-2	Sequence 2, Appl	513	6	1.2	779	1	US-08-426-627-4	Sequence 4, Appl
441	6	1.2	519	1	US-08-589-446-8	Sequence 8, Appl	514	6	1.2	779	1	US-08-426-627-24	Sequence 24, Appl
442	6	1.2	519	2	US-08-444-882-8	Sequence 8, Appl	515	6	1.2	780	1	US-08-375-709-13	Sequence 13, Appl
443	6	1.2	519	2	US-08-389-459A-8	Sequence 8, Appl	516	6	1.2	780	1	US-08-752-929-13	Sequence 13, Appl
444	6	1.2	519	3	US-08-987-867A-8	Sequence 8, Appl	517	6	1.2	788	1	US-08-194-338-12	Sequence 12, Appl
445	6	1.2	524	1	US-08-447-500-24	Sequence 24, Appl	518	6	1.2	788	1	US-08-572-225-1	Sequence 1, Appl
446	6	1.2	524	1	US-08-454-097-24	Sequence 24, Appl	519	6	1.2	811	1	US-08-426-627-2	Sequence 2, Appl
447	6	1.2	524	3	US-08-453-866-24	Sequence 24, Appl	520	6	1.2	811	1	US-08-426-627-22	Sequence 22, Appl
448	6	1.2	524	3	US-08-185-359-24	Sequence 24, Appl	521	6	1.2	816	1	US-07-640-029-1	Sequence 1, Appl
449	6	1.2	525	1	US-08-375-709-19	Sequence 19, Appl	522	6	1.2	817	1	US-07-640-029-2	Sequence 2, Appl
450	6	1.2	525	1	US-08-752-929-19	Sequence 19, Appl	523	6	1.2	820	1	US-07-921-807B-3	Sequence 3, Appl
451	6	1.2	526	2	US-08-852-401-3	Sequence 3, Appl	524	6	1.2	820	1	US-08-291-956-2	Sequence 2, Appl
452	6	1.2	527	1	US-08-160-670A-8	Sequence 8, Appl	525	6	1.2	820	1	US-08-441-944A-3	Sequence 3, Appl
453	6	1.2	530	3	US-09-105-039A-4	Sequence 4, Appl	526	6	1.2	820	1	US-08-166-717D-6	Sequence 6, Appl
454	6	1.2	532	3	US-08-737-336-6	Sequence 6, Appl	527	6	1.2	820	2	US-08-485-278-2	Sequence 2, Appl
455	6	1.2	537	3	US-08-920-610-2	Sequence 2, Appl	528	6	1.2	822	1	US-07-997-133-1	Sequence 1, Appl
456	6	1.2	539	3	US-08-906-360-1	Sequence 1, Appl	529	6	1.2	822	1	US-07-921-807B-4	Sequence 4, Appl
457	6	1.2	547	1	US-08-785-048-3	Sequence 3, Appl	530	6	1.2	822	1	US-08-459-296-2	Sequence 2, Appl
458	6	1.2	547	2	US-08-996-799-3	Sequence 3, Appl	531	6	1.2	822	1	US-08-441-944A-14	Sequence 14, Appl
459	6	1.2	550	4	PCT-US93-03027-1	Sequence 1, Appl	532	6	1.2	822	2	US-08-451-822A-12	Sequence 12, Appl
460	6	1.2	553	1	US-08-785-048-2	Sequence 2, Appl	533	6	1.2	825	1	US-08-642-255-132	Sequence 132, App
461	6	1.2	553	2	US-08-661-052-16	Sequence 16, Appl	534	6	1.2	829	1	US-08-337-633A-53	Sequence 53, Appl
462	6	1.2	553	2	US-08-996-799-2	Sequence 2, Appl	535	6	1.2	832	3	US-08-630-820-7	Sequence 7, Appl
463	6	1.2	559	1	US-08-320-559-31	Sequence 31, Appl	536	6	1.2	836	1	US-08-426-627-6	Sequence 6, Appl
464	6	1.2	559	1	US-08-545-860D-31	Sequence 31, Appl	537	6	1.2	837	1	US-08-426-627-23	Sequence 23, Appl
465	6	1.2	559	4	PCT-US94-04496-31	Sequence 31, Appl	538	6	1.2	856	2	US-07-916-098A-2	Sequence 2, Appl
466	6	1.2	561	1	US-08-360-673-2	Sequence 2, Appl	539	6	1.2	870	2	US-08-732-192A-2	Sequence 2, Appl

248	6	1.2	55	2	US-08-464-134-273	Sequence 273, App	321	6	1.2	223	3	US-08-837-058-9	Sequence 9, Appl
249	6	1.2	55	2	US-08-461-361-273	Sequence 273, App	322	6	1.2	242	3	US-09-167-434-8	Sequence 8, Appl
250	6	1.2	55	2	US-08-485-910-273	Sequence 273, App	323	6	1.2	243	3	US-09-167-434-9	Sequence 9, Appl
251	6	1.2	57	2	US-08-652-816A-41	Sequence 41, Appl	324	6	1.2	243	3	US-08-480-753-4	Sequence 4, Appl
252	6	1.2	57	2	US-08-958-201-48	Sequence 16, Appl	325	6	1.2	246	2	US-09-167-434-11	Sequence 11, Appl
253	6	1.2	67	2	US-08-142-551B-9	Sequence 9, Appl	326	6	1.2	246	2	PCT-US93-03895-8	Sequence 8, Appl
254	6	1.2	69	2	US-08-687-865A-21	Sequence 21, Appl	327	6	1.2	250	4	PCT-US93-03895-6	Sequence 6, Appl
255	6	1.2	71	3	US-08-783-974-6	Sequence 6, Appl	328	6	1.2	251	1	US-08-397-633A-74	Sequence 74, Appl
256	6	1.2	79	2	US-08-448-418-57	Sequence 97, Appl	329	6	1.2	252	1	US-08-792-019B-13	Sequence 13, Appl
257	6	1.2	81	2	US-08-717-159-15	Sequence 15, Appl	330	6	1.2	252	3	US-08-988-819-13	Sequence 13, Appl
258	6	1.2	81	2	US-08-717-159-15	Sequence 15, Appl	331	6	1.2	252	3	US-08-792-824-4	Sequence 4, Appl
259	6	1.2	85	3	US-08-717-159-15	Sequence 32, Appl	332	6	1.2	254	2	US-08-792-824-7	Sequence 7, Appl
260	6	1.2	87	1	US-08-160-670A-7	Sequence 7, Appl	333	6	1.2	254	2	US-08-792-824-10	Sequence 10, Appl
261	6	1.2	89	1	US-08-340-812-13	Sequence 13, Appl	334	6	1.2	254	2	US-08-792-824-13	Sequence 13, Appl
262	6	1.2	89	2	US-08-459-064B-13	Sequence 13, Appl	335	6	1.2	255	3	US-09-167-434-13	Sequence 10, Appl
263	6	1.2	89	2	US-08-460-421A-13	Sequence 13, Appl	336	6	1.2	257	3	US-08-486-099-113	Sequence 113, App
264	6	1.2	89	2	US-08-717-159-7	Sequence 7, Appl	337	6	1.2	257	3	US-08-360-107A-113	Sequence 113, App
265	6	1.2	89	2	US-08-946-026-2	Sequence 2, Appl	338	6	1.2	257	3	US-08-484-223B-113	Sequence 113, App
266	6	1.2	94	1	US-08-690-011A-9	Sequence 9, Appl	339	6	1.2	257	3	US-08-919-597-113	Sequence 113, App
267	6	1.2	94	2	US-08-266-451B-20	Sequence 20, Appl	340	6	1.2	257	3	US-08-475-668A-113	Sequence 113, App
268	6	1.2	94	2	US-08-748-725-50	Sequence 20, Appl	341	6	1.2	257	3	US-08-475-668A-113	Sequence 113, App
269	6	1.2	97	3	US-08-816-977-23	Sequence 23, Appl	342	6	1.2	257	3	US-08-485-551A-113	Sequence 113, App
270	6	1.2	97	3	US-08-816-977-27	Sequence 27, Appl	343	6	1.2	261	1	US-08-160-670A-45	Sequence 45, App
271	6	1.2	98	3	US-08-478-097A-4	Sequence 4, Appl	344	6	1.2	262	1	US-08-397-633A-73	Sequence 73, Appl
272	6	1.2	99	2	US-08-672-345C-94	Sequence 94, Appl	345	6	1.2	263	2	US-08-160-670A-32	Sequence 32, Appl
273	6	1.2	99	3	US-08-484-905-62	Sequence 62, Appl	346	6	1.2	263	2	US-08-353-476-66	Sequence 66, Appl
274	6	1.2	109	2	US-08-481-985B-62	Sequence 62, Appl	347	6	1.2	264	2	US-08-353-476-67	Sequence 67, Appl
275	6	1.2	112	1	US-08-672-345C-4	Sequence 4, Appl	348	6	1.2	264	2	US-08-436-748-8	Sequence 8, Appl
276	6	1.2	112	1	US-08-211-202-133	Sequence 133, App	349	6	1.2	264	2	US-07-857-224B-9	Sequence 9, Appl
277	6	1.2	116	1	US-08-478-039-67	Sequence 67, Appl	350	6	1.2	275	2	US-08-483-857-8	Sequence 8, Appl
278	6	1.2	116	1	US-08-476-349A-67	Sequence 67, Appl	351	6	1.2	281	1	US-08-578-703A-13	Sequence 13, Appl
279	6	1.2	116	1	US-08-481-377-26	Sequence 26, Appl	352	6	1.2	281	1	US-08-397-633A-75	Sequence 75, Appl
280	6	1.2	116	2	US-08-491-835-24	Sequence 24, Appl	353	6	1.2	287	1	US-08-397-633A-76	Sequence 76, Appl
281	6	1.2	116	3	US-09-153-733A-26	Sequence 26, Appl	354	6	1.2	287	1	US-08-011-398B-4	Sequence 4, Appl
282	6	1.2	116	4	PCT-US94-00666-26	Sequence 26, Appl	355	6	1.2	297	2	US-08-464-051-4	Sequence 4, Appl
283	6	1.2	120	4	PCT-US94-00685-24	Sequence 24, Appl	356	6	1.2	297	2	US-08-462-198-4	Sequence 4, Appl
284	6	1.2	121	4	PCT-US93-03895-2	Sequence 2, Appl	357	6	1.2	297	2	US-08-554-485-3	Sequence 3, Appl
285	6	1.2	125	1	US-08-268-348A-12	Sequence 12, Appl	358	6	1.2	304	1	US-09-080-044-9	Sequence 9, Appl
286	6	1.2	127	1	US-08-253-155A-35	Sequence 35, Appl	359	6	1.2	304	1	US-08-661-052-14	Sequence 14, Appl
287	6	1.2	128	1	US-08-426-627-14	Sequence 35, Appl	360	6	1.2	304	1	US-08-132-403-4	Sequence 4, Appl
288	6	1.2	133	1	US-08-268-348A-8	Sequence 8, Appl	361	6	1.2	304	4	US-08-395-939A-4	Sequence 4, Appl
289	6	1.2	133	1	US-08-268-348A-10	Sequence 10, Appl	362	6	1.2	308	1	PCT-US91-01861-4	Sequence 4, Appl
290	6	1.2	172	2	US-08-557-309B-39	Sequence 39, Appl	363	6	1.2	308	1	US-08-499-568-2	Sequence 2, Appl
291	6	1.2	172	3	US-08-834-306-39	Sequence 39, Appl	364	6	1.2	311	3	US-08-793-958-2	Sequence 2, Appl
292	6	1.2	174	3	US-08-906-136A-2	Sequence 2, Appl	365	6	1.2	312	1	US-08-946-214-4	Sequence 4, Appl
293	6	1.2	174	3	US-08-906-136A-4	Sequence 4, Appl	366	6	1.2	315	2	US-08-414-525-2	Sequence 2, Appl
294	6	1.2	179	4	US-08-809-267-7	Sequence 7, Appl	367	6	1.2	315	2	US-08-484-556-91	Sequence 91, Appl
295	6	1.2	179	4	PCT-US95-13662A-7	Sequence 7, Appl	368	6	1.2	319	3	US-08-757-553-91	Sequence 91, Appl
296	6	1.2	180	3	US-08-772-440-31	Sequence 31, Appl	369	6	1.2	320	2	US-08-630-172-12	Sequence 12, Appl
297	6	1.2	183	5	5168049-3	Sequence 31, Appl	370	6	1.2	320	2	US-08-757-553-163	Sequence 163, App
298	6	1.2	184	3	US-09-023-082A-30	Sequence 30, Appl	371	6	1.2	323	3	US-08-823-516-61	Sequence 61, Appl
299	6	1.2	197	1	US-08-879-561-5	Sequence 5, Appl	372	6	1.2	323	3	US-08-816-977-21	Sequence 21, Appl
300	6	1.2	198	1	US-08-642-255-131	Sequence 131, App	373	6	1.2	326	3	US-08-816-977-25	Sequence 25, Appl
301	6	1.2	198	1	US-08-397-633A-41	Sequence 41, Appl	374	6	1.2	338	2	US-08-447-430A-43	Sequence 43, Appl
302	6	1.2	198	1	US-08-397-633A-52	Sequence 52, Appl	375	6	1.2	343	2	US-09-123-851-1	Sequence 1, Appl
303	6	1.2	214	2	US-08-846-790A-1	Sequence 1, Appl	376	6	1.2	343	2	US-08-728-820-1	Sequence 1, Appl
304	6	1.2	214	3	US-08-935-333-1	Sequence 1, Appl	377	6	1.2	344	2	US-07-941-523-24	Sequence 24, Appl
305	6	1.2	218	1	US-08-644-664B-14	Sequence 14, Appl	378	6	1.2	351	2	US-08-701-191A-3	Sequence 3, Appl
306	6	1.2	218	2	US-08-761-277-14	Sequence 14, Appl	379	6	1.2	357	1	US-08-078-683A-8	Sequence 8, Appl
307	6	1.2	219	1	US-08-621-081A-16	Sequence 16, Appl	380	6	1.2	360	1	US-08-674-421-3	Sequence 3, Appl
308	6	1.2	223	1	US-08-430-633-1	Sequence 1, Appl	381	6	1.2	360	1	US-08-469-421-14	Sequence 14, Appl
309	6	1.2	223	1	US-08-936-854-1	Sequence 1, Appl	382	6	1.2	360	1	US-08-250-975-14	Sequence 14, Appl
310	6	1.2	224	2	US-08-360-673-7	Sequence 7, Appl	383	6	1.2	360	2	US-08-920-296-3	Sequence 3, Appl
311	6	1.2	227	1	US-08-414-625-4	Sequence 4, Appl	384	6	1.2	360	2	US-08-605-002A-14	Sequence 14, Appl
312	6	1.2	227	5	5428012-4	Sequence 4, Appl	385	6	1.2	360	2	US-08-950-449A-14	Sequence 14, Appl
313	6	1.2	227	5	5428012-4	Sequence 4, Appl	386	6	1.2	360	3	US-08-554-385-20	Sequence 20, Appl
314	6	1.2	227	5	5451506-3	Sequence 3, App	387	6	1.2	360	4	PCT-US94-10529-14	Sequence 14, Appl
315	6	1.2	228	5	5428012-2	Sequence 2, App	388	6	1.2	376	3	US-08-751-512-8	Sequence 8, Appl
316	6	1.2	228	5	5451506-2	Sequence 2, App	389	6	1.2	377	2	US-08-929-522B-2	Sequence 2, Appl
317	6	1.2	233	1	US-08-446-918A-4	Sequence 4, Appl	390	6	1.2	378	2	US-08-707-337A-104	Sequence 104, App
318	6	1.2	233	2	US-08-480-753-2	Sequence 2, Appl	391	6	1.2	378	4	US-08-642-246-26	Sequence 26, Appl
319	6	1.2	233	2	US-08-580-806-4	Sequence 4, Appl	392	6	1.2	380	1	PCT-US96-06229-26	Sequence 26, Appl
320	6	1.2	233	3	US-09-041-889-9	Sequence 9, Appl	393	6	1.2	380	1	US-08-585-758A-4	Sequence 4, Appl

394	6	1.2	380	2	US-08-670-27AB-4	Sequence 4, Appl1	467	6	1.2	593	1	US-08-374-155A-14	Sequence 14, Appl1
395	6	1.2	381	1	US-08-585-758A-2	Sequence 2, Appl1	468	6	1.2	593	2	US-08-785-356-14	Sequence 14, Appl1
396	6	1.2	381	2	US-08-977-81B-2	Sequence 2, Appl1	469	6	1.2	601	2	US-08-606-288-7	Sequence 7, Appl1
397	6	1.2	381	2	US-08-670-27AB-2	Sequence 2, Appl1	470	6	1.2	601	2	US-08-606-288-10	Sequence 10, Appl1
398	6	1.2	381	2	US-08-786-999-3	Sequence 3, Appl1	471	6	1.2	612	2	US-08-753-307B-11	Sequence 11, Appl1
399	6	1.2	386	3	US-08-895-707-7	Sequence 7, Appl1	472	6	1.2	616	3	US-08-895-707-2	Sequence 2, Appl1
400	6	1.2	390	2	US-08-786-999-1	Sequence 1, Appl1	473	6	1.2	628	3	US-08-394-177-6	Sequence 6, Appl1
401	6	1.2	391	1	US-07-995-657-6	Sequence 6, Appl1	474	6	1.2	630	2	US-08-596-319-2	Sequence 12, Appl1
402	6	1.2	391	1	US-08-474-587-6	Sequence 6, Appl1	475	6	1.2	662	2	US-08-605-541B-12	Sequence 12, Appl1
403	6	1.2	391	2	US-08-244-951A-10	Sequence 10, Appl1	476	6	1.2	662	2	US-08-770-761A-1	Sequence 5, Appl1
404	6	1.2	391	2	US-08-389-011-23	Sequence 23, Appl1	477	6	1.2	677	3	US-08-480-640A-115	Sequence 115, App
405	6	1.2	391	3	US-08-403-917A-23	Sequence 23, Appl1	478	6	1.2	677	3	US-08-480-640A-193	Sequence 193, App
406	6	1.2	391	4	PCR-US85-1566-2	Sequence 2, Appl1	479	6	1.2	680	1	US-08-542-353-4	Sequence 4, Appl1
407	6	1.2	393	4	US-08-577-554-2	Sequence 2, Appl1	480	6	1.2	682	1	US-08-642-255-126	Sequence 126, App
408	6	1.2	401	5	5252556-1	Patent No. 5252556	481	6	1.2	682	3	US-08-481-435-6	Sequence 6, Appl1
409	6	1.2	403	2	US-09-061-337-10	Sequence 10, Appl1	482	6	1.2	705	2	US-08-533-669A-16	Sequence 16, Appl1
410	6	1.2	403	3	US-09-122-129-10	Sequence 10, Appl1	483	6	1.2	705	2	US-08-770-761A-7	Sequence 7, Appl1
411	6	1.2	406	2	US-09-340-991-10	Sequence 10, Appl1	484	6	1.2	729	1	US-07-640-039-3	Sequence 3, Appl1
412	6	1.2	406	2	US-08-222-719-1	Sequence 1, Appl1	485	6	1.2	731	1	US-07-921-807B-5	Sequence 5, Appl1
413	6	1.2	406	2	US-08-470-925-1	Sequence 1, Appl1	486	6	1.2	731	1	US-08-441-944A-5	Sequence 5, Appl1
414	6	1.2	406	2	US-08-471-613-1	Sequence 1, Appl1	487	6	1.2	733	1	US-07-640-029-4	Sequence 4, Appl1
415	6	1.2	407	2	PCR-US93-10443-1	Sequence 1, Appl1	488	6	1.2	733	1	US-07-921-807B-6	Sequence 6, Appl1
416	6	1.2	407	2	US-08-776-585-3	Sequence 3, Appl1	489	6	1.2	733	1	US-08-441-944A-6	Sequence 6, Appl1
417	6	1.2	423	2	US-08-910-798-2	Sequence 2, Appl1	490	6	1.2	733	1	US-08-188-352-16	Sequence 16, Appl1
418	6	1.2	423	2	US-08-313-553-13	Sequence 13, Appl1	491	6	1.2	737	1	US-08-646-715-16	Sequence 16, Appl1
419	6	1.2	423	3	US-08-767-993-13	Sequence 13, Appl1	492	6	1.2	741	2	US-08-849-480A-9	Sequence 9, Appl1
420	6	1.2	427	2	US-08-846-021A-8	Sequence 8, Appl1	493	6	1.2	761	2	US-08-707-237A-84	Sequence 84, Appl1
421	6	1.2	439	2	US-08-507-431-4	Sequence 4, Appl1	494	6	1.2	762	1	US-08-642-255-114	Sequence 114, App
422	6	1.2	439	2	US-08-700-546-2	Sequence 2, Appl1	495	6	1.2	762	1	US-08-642-255-120	Sequence 120, App
423	6	1.2	439	2	US-08-902-655A-4	Sequence 4, Appl1	496	6	1.2	762	1	US-08-397-633A-26	Sequence 26, Appl1
424	6	1.2	439	3	US-09-116-622-4	Sequence 4, Appl1	497	6	1.2	762	1	US-08-397-633A-31	Sequence 31, Appl1
425	6	1.2	450	2	US-08-665-037-2	Sequence 2, Appl1	498	6	1.2	764	1	US-08-424-567-2	Sequence 2, Appl1
426	6	1.2	450	2	US-08-666-067-2	Sequence 2, Appl1	499	6	1.2	764	1	US-08-711-928-2	Sequence 2, Appl1
427	6	1.2	450	2	US-08-732-870-2	Sequence 2, Appl1	500	6	1.2	768	2	US-08-560-398-2	Sequence 2, Appl1
428	6	1.2	458	5	5217891-15	Patent No. 5217891	501	6	1.2	770	1	US-08-359-706-12	Sequence 12, Appl1
429	6	1.2	474	2	US-09-096-982-8	Sequence 8, Appl1	502	6	1.2	770	1	US-08-416-581B-1	Sequence 1, Appl1
430	6	1.2	474	2	US-08-653-650A-8	Sequence 8, Appl1	503	6	1.2	770	1	US-08-416-581B-5	Sequence 5, Appl1
431	6	1.2	479	1	US-08-397-633A-78	Sequence 78, Appl1	504	6	1.2	770	1	US-08-416-581B-9	Sequence 9, Appl1
432	6	1.2	481	2	US-08-477-451-19	Sequence 19, Appl1	505	6	1.2	770	2	US-08-853-091-12	Sequence 12, Appl1
433	6	1.2	486	1	US-08-397-633A-77	Sequence 77, Appl1	506	6	1.2	770	2	US-08-820-754-12	Sequence 12, Appl1
434	6	1.2	494	4	US-08-464-340A-4	Sequence 4, Appl1	507	6	1.2	770	3	US-08-956-652-12	Sequence 12, Appl1
435	6	1.2	494	4	PCR-US94-08449A-4	Sequence 4, Appl1	508	6	1.2	770	4	PCR-US95-17025-12	Sequence 12, Appl1
436	6	1.2	498	2	US-07-816-098A-4	Sequence 4, Appl1	509	6	1.2	770	4	US-08-276-099A-14	Sequence 14, Appl1
437	6	1.2	500	2	US-08-578-709-15	Sequence 15, Appl1	510	6	1.2	771	1	US-08-781-890-14	Sequence 14, Appl1
438	6	1.2	504	2	US-08-896-005-3	Sequence 3, Appl1	511	6	1.2	771	1	US-08-426-637-4	Sequence 4, Appl1
439	6	1.2	511	3	US-09-105-039A-2	Sequence 2, Appl1	512	6	1.2	777	2	US-08-426-637-2	Sequence 2, Appl1
440	6	1.2	514	3	US-08-865-311-2	Sequence 8, Appl1	513	6	1.2	777	2	US-08-426-637-2	Sequence 2, Appl1
441	6	1.2	519	1	US-08-589-446-8	Sequence 8, Appl1	514	6	1.2	779	1	US-08-426-637-2	Sequence 2, Appl1
442	6	1.2	519	1	US-08-444-882-8	Sequence 8, Appl1	515	6	1.2	780	1	US-08-375-709-13	Sequence 13, Appl1
443	6	1.2	519	2	US-08-389-455A-8	Sequence 8, Appl1	516	6	1.2	780	1	US-08-752-929-13	Sequence 13, Appl1
444	6	1.2	519	3	US-08-987-867A-8	Sequence 8, Appl1	517	6	1.2	788	1	US-08-194-338-12	Sequence 12, Appl1
445	6	1.2	524	1	US-08-447-500-24	Sequence 24, Appl1	518	6	1.2	788	1	US-08-572-225-1	Sequence 1, Appl1
446	6	1.2	524	1	US-08-454-097-24	Sequence 24, Appl1	519	6	1.2	811	1	US-08-426-637-2	Sequence 2, Appl1
447	6	1.2	524	1	US-08-453-866-24	Sequence 24, Appl1	520	6	1.2	811	1	US-08-426-637-2	Sequence 2, Appl1
448	6	1.2	525	3	US-08-185-359-24	Sequence 24, Appl1	521	6	1.2	816	1	US-07-640-029-1	Sequence 1, Appl1
449	6	1.2	525	1	US-08-375-709-19	Sequence 19, Appl1	522	6	1.2	817	1	US-07-640-029-2	Sequence 2, Appl1
450	6	1.2	525	1	US-08-752-929-19	Sequence 19, Appl1	523	6	1.2	820	1	US-07-921-807B-3	Sequence 3, Appl1
451	6	1.2	526	2	US-08-852-401-3	Sequence 3, Appl1	524	6	1.2	820	1	US-08-291-896-2	Sequence 2, Appl1
452	6	1.2	527	2	US-08-160-670A-8	Sequence 8, Appl1	525	6	1.2	820	1	US-08-441-944A-3	Sequence 3, Appl1
453	6	1.2	530	3	US-09-105-039A-4	Sequence 4, Appl1	526	6	1.2	820	1	US-08-166-717D-6	Sequence 6, Appl1
454	6	1.2	530	3	US-08-737-336-6	Sequence 6, Appl1	527	6	1.2	820	2	US-08-485-278-2	Sequence 2, Appl1
455	6	1.2	537	3	US-08-920-610-2	Sequence 2, Appl1	528	6	1.2	822	1	US-07-997-807B-4	Sequence 4, Appl1
456	6	1.2	539	3	US-08-906-360-1	Sequence 1, Appl1	529	6	1.2	822	1	US-07-921-807B-4	Sequence 4, Appl1
457	6	1.2	547	1	US-08-785-048-3	Sequence 3, Appl1	530	6	1.2	822	1	US-08-459-286-2	Sequence 2, Appl1
458	6	1.2	547	1	US-08-996-799-3	Sequence 3, Appl1	531	6	1.2	822	1	US-08-441-944A-4	Sequence 4, Appl1
459	6	1.2	550	4	PCR-US93-03027-1	Sequence 1, Appl1	532	6	1.2	822	2	US-08-451-822A-12	Sequence 12, Appl1
460	6	1.2	553	1	US-08-785-048-2	Sequence 2, Appl1	533	6	1.2	829	1	US-08-642-255-132	Sequence 132, App
461	6	1.2	553	2	US-08-661-052-16	Sequence 16, Appl1	534	6	1.2	829	3	US-08-397-633A-53	Sequence 53, Appl1
462	6	1.2	553	2	US-08-996-799-2	Sequence 2, Appl1	535	6	1.2	836	1	US-08-630-820-7	Sequence 7, Appl1
463	6	1.2	559	1	US-08-320-559-31	Sequence 31, Appl1	536	6	1.2	837	1	US-08-426-637-6	Sequence 6, Appl1
464	6	1.2	559	3	US-08-545-860D-31	Sequence 31, Appl1	537	6	1.2	836	2	US-07-916-098A-2	Sequence 2, Appl1
465	6	1.2	559	4	PCR-US94-04496-31	Sequence 31, Appl1	538	6	1.2	870	2	US-08-732-192A-2	Sequence 2, Appl1
466	6	1.2	561	1	US-08-360-673-2	Sequence 2, Appl1	539	6	1.2				

540	6	1.2	872	2	US-08-844-057-2	Sequence 2, Appl1	613	6	1.2	2105	2	US-08-808-793-3	Sequence 3, Appl1
541	6	1.2	877	1	US-08-397-633a-54	Sequence 54, Appl1	614	6	1.2	2105	3	US-08-772-512A-3	Sequence 3, Appl1
542	6	1.2	878	1	US-08-37-919-2	Sequence 2, Appl1	615	6	1.2	2206	1	US-07-852-260-2	Sequence 2, Appl1
543	6	1.2	878	4	PCT-US95-0518-2	Sequence 2, Appl1	616	6	1.2	2206	1	US-08-461-503-2	Sequence 2, Appl1
544	6	1.2	884	2	US-08-397-633a-68	Sequence 68, Appl1	617	6	1.2	2882	1	US-07-741-940-7	Sequence 7, Appl1
545	6	1.2	884	2	US-08-474-067-8	Sequence 8, Appl1	618	6	1.2	2882	1	US-08-889-548A-7	Sequence 7, Appl1
546	6	1.2	884	2	US-08-435-641-15	Sequence 15, Appl1	619	6	1.2	2882	1	US-08-452-654-7	Sequence 7, Appl1
547	6	1.2	884	2	US-08-707-237a-96	Sequence 96, Appl1	620	6	1.2	2843	1	US-07-741-940-2	Sequence 2, Appl1
548	6	1.2	884	2	US-08-474-068a-8	Sequence 8, Appl1	621	6	1.2	2843	1	US-08-289-548A-2	Sequence 2, Appl1
549	6	1.2	884	2	US-08-472-481-7	Sequence 7, Appl1	622	6	1.2	2843	1	US-08-452-654-2	Sequence 2, Appl1
550	6	1.2	884	3	US-08-642-246-15	Sequence 15, Appl1	623	6	1.2	2843	1	US-08-452-655B-7	Sequence 7, Appl1
551	6	1.2	884	4	PCT-US96-0622a-15	Sequence 15, Appl1	624	6	1.2	2843	1	US-08-452-655B-7	Sequence 7, Appl1
552	6	1.2	933	2	US-08-313-200-1	Sequence 1, Appl1	625	6	1.2	2843	2	US-08-370-235A-2	Sequence 2, Appl1
553	6	1.2	933	4	PCT-US93-03837-1	Sequence 1, Appl1	626	6	1.2	2973	2	US-08-821-355A-7	Sequence 7, Appl1
554	6	1.2	936	2	US-08-707-237a-108	Sequence 108, App	627	6	1.2	2973	2	US-09-003-687A-7	Sequence 7, Appl1
555	6	1.2	936	3	US-08-642-246-30	Sequence 30, Appl1	628	6	1.2	3391	2	US-07-642-734C-2	Sequence 2, Appl1
556	6	1.2	936	4	PCT-US96-0622a-30	Sequence 30, Appl1	629	6	1.2	3391	2	US-07-642-734C-4	Sequence 4, Appl1
557	6	1.2	966	3	US-08-642-246-34	Sequence 34, Appl1	630	6	1.2	3391	2	US-08-804-227C-10	Sequence 10, Appl1
558	6	1.2	966	4	PCT-US96-0622a-34	Sequence 34, Appl1	631	6	1.2	3724	2	US-08-804-227C-10	Sequence 4, Appl1
559	6	1.2	970	2	US-08-449-645a-11	Sequence 11, Appl1	632	6	1.2	3724	2	US-08-471-119A-2	Sequence 2, Appl1
560	6	1.2	970	2	US-08-702-367a-11	Sequence 11, Appl1	633	6	1.2	15381	1	US-07-816-678A-7	Sequence 7, Appl1
561	6	1.2	980	4	PCT-US95-04681-11	Sequence 11, Appl1	634	6	1.0	5	1	US-07-826-928A-7	Sequence 7, Appl1
562	6	1.2	980	1	US-08-091-569-2	Sequence 2, Appl1	635	6	1.0	5	1	US-08-197-496A-21	Sequence 21, Appl1
563	6	1.2	980	1	US-08-220-151-5	Sequence 5, Appl1	636	6	1.0	5	1	US-08-131-365B-49	Sequence 49, Appl1
564	6	1.2	980	1	US-08-203-676-2	Sequence 2, Appl1	637	6	1.0	5	1	US-08-275-370-26	Sequence 26, Appl1
565	6	1.2	980	1	US-08-413-118-5	Sequence 5, Appl1	638	6	1.0	5	1	US-08-367-968-26	Sequence 26, Appl1
566	6	1.2	980	2	US-08-822-238-2	Sequence 2, Appl1	639	6	1.0	5	1	US-08-294-434-2	Sequence 2, Appl1
567	6	1.2	980	3	US-08-473-446-5	Sequence 5, Appl1	640	6	1.0	5	1	US-08-457-166-2	Sequence 2, Appl1
568	6	1.2	990	2	US-08-392-625-20	Sequence 20, Appl1	641	6	1.0	5	1	US-08-200-900A-34	Sequence 34, Appl1
569	6	1.2	990	2	US-08-466-961a-20	Sequence 20, Appl1	642	6	1.0	5	1	US-08-480-604A-32	Sequence 32, Appl1
570	6	1.2	990	2	US-08-545-193a-15	Sequence 15, Appl1	643	6	1.0	5	1	US-08-665-484-26	Sequence 26, Appl1
571	6	1.2	995	1	US-08-162-809-18	Sequence 18, Appl1	644	6	1.0	5	2	US-07-963-538B-11	Sequence 11, Appl1
572	6	1.2	995	2	US-08-673-789-5	Sequence 5, Appl1	645	6	1.0	5	2	US-08-668-122-45	Sequence 45, Appl1
573	6	1.2	995	2	US-08-717-515-8	Sequence 8, Appl1	646	6	1.0	5	2	US-08-605-496A-30	Sequence 30, Appl1
574	6	1.2	995	4	PCT-US95-04910-14	Sequence 14, Appl1	647	6	1.0	5	2	US-09-016-366A-29	Sequence 29, Appl1
575	6	1.2	1002	2	US-08-707-237a-103	Sequence 103, App	648	6	1.0	5	2	US-08-481-433-41	Sequence 41, Appl1
576	6	1.2	1002	3	US-08-642-246-25	Sequence 25, Appl1	649	6	1.0	5	2	US-08-997-918-56	Sequence 56, Appl1
577	6	1.2	1002	4	PCT-US96-0622a-25	Sequence 25, Appl1	650	6	1.0	5	2	PCT-US93-06591-2	Sequence 2, Appl1
578	6	1.2	1011	1	US-08-162-809-12	Sequence 12, Appl1	651	6	1.0	5	3	US-08-483-749A-33	Sequence 33, Appl1
579	6	1.2	1046	1	US-08-386-727-2	Sequence 2, Appl1	652	6	1.0	5	3	US-08-997-918-56	Sequence 56, Appl1
580	6	1.2	1084	1	US-08-600-452a-2	Sequence 2, Appl1	653	6	1.0	5	3	US-08-595-586B-16	Sequence 16, Appl1
581	6	1.2	1194	1	US-08-717-515-6	Sequence 6, Appl1	654	6	1.0	5	3	US-08-483-749A-33	Sequence 33, Appl1
582	6	1.2	1220	2	US-08-680-326-35	Sequence 35, Appl1	655	6	1.0	5	4	PCT-US92-11270-7	Sequence 7, Appl1
583	6	1.2	1220	2	US-08-80-326-38	Sequence 38, Appl1	656	6	1.0	5	4	US-08-955-15800-2	Sequence 2, Appl1
584	6	1.2	1253	1	US-08-252-966B-12	Sequence 12, Appl1	657	6	1.0	5	4	US-08-243-082-21	Sequence 21, Appl1
585	6	1.2	1261	1	US-08-252-966B-18	Sequence 18, Appl1	658	6	1.0	5	4	US-08-343-943-2	Sequence 2, Appl1
586	6	1.2	1276	1	US-08-717-515-8	Sequence 8, Appl1	659	6	1.0	5	4	US-08-481-433-41	Sequence 41, Appl1
587	6	1.2	1394	4	PCT-US95-10661A-2	Sequence 2, Appl1	660	6	1.0	5	4	US-08-243-082-21	Sequence 2, Appl1
588	6	1.2	1398	1	US-08-750-532-9	Sequence 9, Appl1	661	6	1.0	5	4	US-08-115-519-6	Sequence 6, Appl1
589	6	1.2	1452	2	US-08-652-971-4	Sequence 4, Appl1	662	6	1.0	5	4	US-08-200-900A-35	Sequence 35, Appl1
590	6	1.2	1452	2	US-08-991-258A-4	Sequence 4, Appl1	663	6	1.0	5	4	US-07-885-089B-19	Sequence 19, Appl1
591	6	1.2	1452	2	US-08-769-399-4	Sequence 4, Appl1	664	6	1.0	5	4	US-08-531-525-1	Sequence 1, Appl1
592	6	1.2	1477	1	US-08-038-682-4	Sequence 4, Appl1	665	6	1.0	5	4	US-07-963-538B-6	Sequence 6, Appl1
593	6	1.2	1477	1	US-08-302-832-4	Sequence 4, Appl1	666	6	1.0	5	4	US-08-692-825-39	Sequence 39, Appl1
594	6	1.2	1477	2	US-08-530-198-4	Sequence 4, Appl1	667	6	1.0	5	4	US-08-718-270A-1	Sequence 1, Appl1
595	6	1.2	1477	2	US-08-469-880-4	Sequence 4, Appl1	668	6	1.0	5	4	US-08-945-742-8	Sequence 8, Appl1
596	6	1.2	1477	2	US-08-728-470-4	Sequence 4, Appl1	669	6	1.0	5	4	US-08-878-404B-24	Sequence 24, Appl1
597	6	1.2	1477	2	US-08-617-697-4	Sequence 4, Appl1	670	6	1.0	5	4	US-08-951-648-34	Sequence 34, Appl1
598	6	1.2	1541	4	PCT-US95-10661A-3	Sequence 4, Appl1	671	6	1.0	5	4	US-08-481-433-28	Sequence 28, Appl1
599	6	1.2	1566	2	US-08-687-956A-23	Sequence 23, Appl1	672	6	1.0	5	4	US-08-704-170-43	Sequence 43, Appl1
600	6	1.2	1620	1	US-08-542-363-2	Sequence 2, Appl1	673	6	1.0	5	4	US-08-899-575-14	Sequence 14, Appl1
601	6	1.2	1621	1	US-08-242-677-2	Sequence 2, Appl1	674	6	1.0	5	4	US-08-115-519-7	Sequence 7, Appl1
602	6	1.2	1621	1	US-07-665-792E-9	Sequence 9, Appl1	675	6	1.0	5	4	US-08-951-648-34	Sequence 34, Appl1
603	6	1.2	1648	2	US-08-662-227-35	Sequence 35, Appl1	676	6	1.0	5	4	US-08-481-433-28	Sequence 28, Appl1
604	6	1.2	1664	2	US-08-642-846-2	Sequence 2, Appl1	677	6	1.0	5	4	PCT-US94-02631-43	Sequence 43, Appl1
605	6	1.2	1820	2	US-07-998-289E-8	Sequence 8, Appl1	678	6	1.0	5	4	PCT-US95-08743-14	Sequence 14, Appl1
606	6	1.2	1820	2	US-08-804-227C-6	Sequence 6, Appl1	679	6	1.0	5	4	US-08-323-474-4	Sequence 4, Appl1
607	6	1.2	1958	1	US-07-945-283-2	Sequence 2, Appl1	680	6	1.0	5	4	US-08-825-989-15	Sequence 15, Appl1
608	6	1.2	2089	1	US-08-418-893D-23	Sequence 23, Appl1	681	6	1.0	5	4	US-08-086-335C-12	Sequence 12, Appl1
609	6	1.2	2089	1	US-08-418-893D-24	Sequence 24, Appl1	682	6	1.0	5	4	US-07-982-174-1	Sequence 1, Appl1
610	6	1.2	2100	3	US-08-808-793-23	Sequence 23, Appl1	683	6	1.0	5	4	US-08-263-086-5	Sequence 5, Appl1
611	6	1.2	2100	3	US-08-772-512A-19	Sequence 19, Appl1	684	6	1.0	5	4	US-08-318-193-30	Sequence 30, Appl1
612	6	1.2	2104	2	US-08-808-793-4	Sequence 4, Appl1	685	6	1.0	5	4	US-08-240-514-34	Sequence 34, Appl1
	6	1.2	2104	3	US-08-772-512A-4	Sequence 4, Appl1		6	1.0	5	4	US-08-236-918A-16	Sequence 16, Appl1

666	5	1.0	8	1	US-08-570-923-15	Sequence 15, Appl	759	1.0	8	4	PCT-US95-00362-3	Sequence 3, Appl
667	5	1.0	8	1	US-08-446-908-6	Sequence 6, Appl	760	1.0	8	4	PCT-US95-06530-7	Sequence 7, Appl
668	5	1.0	8	1	US-08-297-706-1	Sequence 1, Appl	761	1.0	8	4	PCT-US95-08534-3	Sequence 3, Appl
669	5	1.0	8	1	US-08-634-060-25	Sequence 25, Appl	762	1.0	8	4	PCT-US95-10973A-98	Sequence 98, Appl
670	5	1.0	8	1	US-08-231-205A-6	Sequence 6, Appl	763	1.0	8	4	PCT-US95-12960-2	Sequence 2, Appl
681	5	1.0	8	1	US-08-446-922-1	Sequence 1, Appl	764	1.0	8	4	PCT-US95-15780-6	Sequence 6, Appl
682	5	1.0	8	1	US-08-430-633-3	Sequence 3, Appl	765	1.0	8	4	PCT-US95-15801-19	Sequence 19, Appl
693	5	1.0	8	1	US-08-580-014-15	Sequence 15, Appl	766	1.0	8	4	PCT-US95-15801-19	Sequence 19, Appl
694	5	1.0	8	1	US-08-461-597-8	Sequence 8, Appl	767	1.0	8	4	PCT-US95-15801-19	Sequence 19, Appl
695	5	1.0	8	1	US-08-442-043A-15	Sequence 15, Appl	768	1.0	8	5	5194375-7	Patent No. 5194375
696	5	1.0	8	1	US-08-604-333-5	Sequence 5, Appl	769	1.0	8	5	5198342-3	Patent No. 5198342
697	5	1.0	8	1	US-08-308-881-7	Sequence 7, Appl	770	1.0	9	1	US-08-767-547-98	Sequence 98, Appl
698	5	1.0	8	1	US-08-660-626-1	Sequence 1, Appl	771	1.0	9	2	US-08-340-283-134	Sequence 134, Appl
699	5	1.0	8	1	US-08-526-014-1	Sequence 1, Appl	772	1.0	9	2	US-08-605-002A-17	Sequence 17, Appl
701	5	1.0	8	2	US-08-612-302A-34	Sequence 34, Appl	773	1.0	9	2	US-08-595-043A-41	Sequence 41, Appl
702	5	1.0	8	2	US-08-589-528-1	Sequence 1, Appl	774	1.0	9	2	US-08-859-931A-3	Sequence 3, Appl
703	5	1.0	8	2	US-08-684-687-6	Sequence 6, Appl	775	1.0	9	2	US-08-950-449A-19	Sequence 19, Appl
704	5	1.0	8	2	US-08-701-124-2	Sequence 2, Appl	776	1.0	9	2	US-08-966-234-46	Sequence 46, Appl
705	5	1.0	8	2	US-08-649-341A-9	Sequence 9, Appl	777	1.0	9	3	US-08-968-747-4	Sequence 4, Appl
706	5	1.0	8	2	US-08-494-440B-13	Sequence 13, Appl	778	1.0	9	3	US-08-997-918-53	Sequence 53, Appl
707	5	1.0	8	2	US-08-729-345-4	Sequence 4, Appl	779	1.0	10	1	US-07-721-761A-16	Sequence 16, Appl
708	5	1.0	8	2	US-08-533-901B-15	Sequence 15, Appl	780	1.0	10	1	US-07-978-687-16	Sequence 16, Appl
709	5	1.0	8	2	US-08-670-175-1	Sequence 1, Appl	781	1.0	10	1	US-08-435-201-26	Sequence 26, Appl
710	5	1.0	8	2	US-08-777-405A-15	Sequence 15, Appl	782	1.0	10	1	US-08-568-072-1	Sequence 1, Appl
711	5	1.0	8	2	US-08-922-267A-63	Sequence 63, Appl	783	1.0	10	1	US-08-603-753D-13	Sequence 13, Appl
712	5	1.0	8	2	US-08-535-298-10	Sequence 10, Appl	784	1.0	10	2	US-08-603-753D-14	Sequence 14, Appl
713	5	1.0	8	2	US-08-620-694A-3	Sequence 3, Appl	785	1.0	10	2	US-08-603-753D-15	Sequence 15, Appl
714	5	1.0	8	2	US-08-720-258-7	Sequence 7, Appl	786	1.0	10	2	US-08-603-753D-16	Sequence 16, Appl
715	5	1.0	8	2	US-08-977-871A-15	Sequence 15, Appl	787	1.0	10	2	US-08-603-753D-16	Sequence 16, Appl
716	5	1.0	8	2	US-08-839-032A-19	Sequence 19, Appl	788	1.0	10	2	US-08-713-928B-14	Sequence 14, Appl
717	5	1.0	8	2	US-08-870-518-27	Sequence 27, Appl	789	1.0	10	2	US-08-595-688C-49	Sequence 49, Appl
718	5	1.0	8	2	US-08-936-854-3	Sequence 3, Appl	790	1.0	10	3	US-08-467-580-39	Sequence 39, Appl
719	5	1.0	8	2	US-08-059-099-7	Sequence 7, Appl	791	1.0	10	3	US-08-159-339A-295	Sequence 295, Appl
720	5	1.0	8	2	US-08-713-928B-10	Sequence 10, Appl	792	1.0	10	4	PCT-US91-05801-16	Sequence 16, Appl
721	5	1.0	8	2	US-08-690-011A-50	Sequence 50, Appl	793	1.0	10	4	PCT-US93-04717-26	Sequence 26, Appl
722	5	1.0	8	2	US-08-839-031A-15	Sequence 15, Appl	794	1.0	10	4	PCT-US95-08516-39	Sequence 39, Appl
723	5	1.0	8	2	US-08-762-106-3	Sequence 3, Appl	795	1.0	10	4	PCT-US95-12960-1	Sequence 1, Appl
724	5	1.0	8	2	US-09-016-366A-30	Sequence 30, Appl	796	1.0	11	1	US-08-061-350-2	Sequence 2, Appl
725	5	1.0	8	2	US-08-700-846-8	Sequence 8, Appl	797	1.0	11	1	US-08-200-000A-37	Sequence 37, Appl
726	5	1.0	8	2	US-08-832-935-1	Sequence 5, Appl	798	1.0	11	1	US-08-338-634-34	Sequence 34, Appl
727	5	1.0	8	2	US-08-817-161-6	Sequence 6, Appl	799	1.0	11	1	US-08-471-780C-120	Sequence 120, Appl
728	5	1.0	8	2	US-08-832-935-1	Sequence 1, Appl	800	1.0	11	1	US-08-471-780C-120	Sequence 120, Appl
729	5	1.0	8	2	US-08-482-728A-2	Sequence 2, Appl	801	1.0	11	1	US-08-467-282B-120	Sequence 120, Appl
730	5	1.0	8	2	US-08-825-951-15	Sequence 15, Appl	802	1.0	11	1	US-08-467-282B-120	Sequence 120, Appl
731	5	1.0	8	2	US-09-058-264-7	Sequence 7, Appl	803	1.0	11	1	US-08-467-282B-120	Sequence 120, Appl
732	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	804	1.0	11	1	US-08-471-782A-120	Sequence 120, Appl
733	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	805	1.0	11	1	US-08-471-782A-120	Sequence 120, Appl
734	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	806	1.0	11	2	US-08-318-157B-40	Sequence 40, Appl
735	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	807	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
736	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	808	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
737	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	809	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
738	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	810	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
739	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	811	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
740	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	812	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
741	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	813	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
742	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	814	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
743	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	815	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
744	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	816	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
745	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	817	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
746	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	818	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
747	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	819	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
748	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	820	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
749	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	821	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
750	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	822	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
751	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	823	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
752	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	824	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
753	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	825	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
754	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	826	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
755	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	827	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
756	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	828	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
757	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	829	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
758	5	1.0	8	2	US-08-991-426-8	Sequence 8, Appl	830	1.0	11	2	US-08-466-710C-120	Sequence 120, Appl
							831	1.0	13	2	US-08-860-174A-30	Sequence 30, Appl

832	5	1.0	13	4	PCT-US95-04228-41
833	5	1.0	13	5	5189147-14
834	5	1.0	13	5	5395760-16
835	5	1.0	14	1	US-08-411-795B-403
836	5	1.0	14	1	US-08-446-908-9
837	5	1.0	14	1	US-08-231-205A-9
838	5	1.0	14	2	US-08-469-319A-403
839	5	1.0	14	2	US-08-871-161-9
840	5	1.0	15	1	US-08-190-802A-17
841	5	1.0	15	1	US-08-218-025A-144
842	5	1.0	15	1	US-08-467-083-32
843	5	1.0	15	1	US-08-414-417B-32
844	5	1.0	15	2	US-08-531-525-7
845	5	1.0	15	2	US-08-486-348A-32
846	5	1.0	15	2	US-08-765-179B-8
847	5	1.0	15	2	US-08-468-545B-32
848	5	1.0	15	2	US-08-637-759B-45
849	5	1.0	15	2	US-08-718-270A-7
850	5	1.0	15	2	US-08-310-912A-38
851	5	1.0	15	2	US-08-675-631-4
852	5	1.0	15	2	US-08-553-257A-57
853	5	1.0	15	3	US-08-871-355A-45
854	5	1.0	15	3	US-08-296-848A-5
855	5	1.0	15	3	US-08-660-347-3
856	5	1.0	15	3	US-08-466-680B-32
857	5	1.0	15	4	PCT-US95-04570-38
858	5	1.0	15	4	PCT-US95-04589-38
859	5	1.0	15	5	5189147-20
860	5	1.0	16	1	US-07-789-179-2
861	5	1.0	16	1	US-08-240-712-9
862	5	1.0	16	1	US-08-443-890-9
863	5	1.0	16	1	US-08-444-942-2
864	5	1.0	16	1	US-08-446-105-2
865	5	1.0	16	1	US-08-444-939-2
866	5	1.0	16	2	US-08-482-142-185
867	5	1.0	16	2	US-08-444-991-2
868	5	1.0	16	2	US-08-450-733-2
869	5	1.0	16	2	US-08-478-572-185
870	5	1.0	16	3	US-08-804-439A-100
871	5	1.0	16	3	US-08-720-229-100
872	5	1.0	16	3	US-08-817-145-19
873	5	1.0	16	4	PCT-US91-02942-91
874	5	1.0	16	4	PCT-US92-09752-9
875	5	1.0	16	5	5304631-1
876	5	1.0	17	1	US-08-318-200-24
877	5	1.0	17	1	US-08-257-528B-70
878	5	1.0	17	1	US-08-460-602A-70
879	5	1.0	17	1	US-08-463-966A-70
880	5	1.0	17	1	US-08-465-217A-70
881	5	1.0	17	2	US-08-464-329A-70
882	5	1.0	17	2	US-08-572-951-30
883	5	1.0	17	2	US-08-659-567-20
884	5	1.0	17	2	US-07-963-538B-21
885	5	1.0	17	2	US-08-462-507A-70
886	5	1.0	17	2	US-08-651-818A-22
887	5	1.0	17	2	US-08-467-881A-70
888	5	1.0	18	1	US-07-624-730A-1
889	5	1.0	18	1	US-08-486-135-5
890	5	1.0	18	1	US-08-468-975-2
891	5	1.0	18	1	US-08-470-152-5
892	5	1.0	18	1	US-08-787-547-14
893	5	1.0	18	1	US-08-024-253-14
894	5	1.0	18	2	US-08-468-964B-3
895	5	1.0	18	2	US-09-013-634-4
896	5	1.0	18	2	US-07-871-282A-3
897	5	1.0	18	2	US-08-553-678A-3
898	5	1.0	18	3	US-09-276-202-12
899	5	1.0	18	3	US-08-779-764A-37
900	5	1.0	18	3	US-08-475-577-1
901	5	1.0	18	3	US-08-475-577-3
902	5	1.0	18	3	US-08-475-577-5
903	5	1.0	18	3	US-08-475-577-5
904	5	1.0	18	3	US-08-582-134B-3

Sequence 41, App1	905	5	1.0	18	4	PCT-US91-06532-4
Patent No. 5189147	906	5	1.0	18	4	PCT-US93-07653-1
Patent No. 5395760	907	5	1.0	18	5	5443956-13
Sequence 403, App	908	5	1.0	19	1	US-08-237-418-22
Sequence 9, App1	909	5	1.0	19	5	5304631-10
Sequence 9, App1	910	5	1.0	20	1	US-07-987-286-18
Sequence 403, App	911	5	1.0	20	1	US-08-440-861-39
Sequence 9, App1	912	5	1.0	20	1	US-08-614-935-19
Sequence 17, App1	913	5	1.0	20	2	US-08-614-628-18
Sequence 144, App	914	5	1.0	20	2	US-08-934-915-36
Sequence 32, App1	915	5	1.0	20	2	US-08-934-915-43
Sequence 32, App1	916	5	1.0	20	2	US-08-934-915-81
Sequence 7, App1	917	5	1.0	20	2	US-08-934-915-88
Sequence 32, App1	918	5	1.0	20	2	US-08-934-915-143
Sequence 8, App1	919	5	1.0	20	2	US-08-802-322-8
Sequence 32, App1	920	5	1.0	20	2	US-08-053-451B-122
Sequence 45, App1	921	5	1.0	20	5	5262298-1
Sequence 7, App1	922	5	1.0	21	1	US-07-842-089E-21
Sequence 38, App1	923	5	1.0	21	1	US-07-927-071-3
Sequence 4, App1	924	5	1.0	21	1	US-08-264-485-21
Sequence 57, App1	925	5	1.0	21	1	US-08-471-780C-125
Sequence 45, App1	926	5	1.0	21	1	US-08-467-282B-125
Sequence 5, App1	927	5	1.0	21	1	US-08-471-282A-125
Sequence 3, App1	928	5	1.0	21	2	US-08-466-710C-125
Sequence 32, App1	929	5	1.0	21	3	US-08-468-739C-125
Sequence 38, App1	930	5	1.0	21	5	5242798-16
Sequence 38, App1	931	5	1.0	22	2	US-08-701-124-29
Patent No. 5189147	932	5	1.0	22	2	US-08-802-322-9
Sequence 2, App1	933	5	1.0	22	2	US-08-802-322-11
Sequence 9, App1	934	5	1.0	22	2	US-08-969-721-10
Sequence 9, App1	935	5	1.0	22	2	US-08-378-548-20
Sequence 2, App1	936	5	1.0	22	3	US-08-256-747C-10
Sequence 2, App1	937	5	1.0	22	3	US-08-256-747C-41
Sequence 185, App	938	5	1.0	22	3	US-09-276-202-9
Sequence 2, App1	939	5	1.0	22	4	US-09-130-225-29
Sequence 2, App1	940	5	1.0	22	4	PCT-US91-02942-100
Sequence 185, App	941	5	1.0	23	1	US-08-346-293-21
Sequence 2, App1	942	5	1.0	23	1	US-08-086-428B-156
Sequence 100, App	943	5	1.0	23	1	US-08-447-411-68
Sequence 100, App	944	5	1.0	23	2	US-08-482-142-182
Sequence 19, App1	945	5	1.0	23	2	US-08-493-235-26
Sequence 91, App1	946	5	1.0	23	2	US-08-493-235-27
Sequence 9, App1	947	5	1.0	23	2	US-08-468-570-156
Patent No. 5304631	948	5	1.0	23	2	US-08-290-665A-260
Sequence 24, App1	949	5	1.0	23	2	US-08-662-227-26
Sequence 70, App1	950	5	1.0	23	2	US-08-478-572-182
Sequence 70, App1	951	5	1.0	23	3	US-09-276-202-11
Sequence 70, App1	952	5	1.0	23	4	PCT-US95-10398-260
Sequence 70, App1	953	5	1.0	24	1	US-08-133-011-9
Sequence 70, App1	954	5	1.0	24	1	US-08-322-730A-9
Sequence 30, App1	955	5	1.0	24	1	US-08-387-87A-9
Sequence 20, App1	956	5	1.0	24	2	US-08-637-759B-113
Sequence 21, App1	957	5	1.0	24	2	US-08-383-619-9
Sequence 70, App1	958	5	1.0	24	3	US-08-871-355A-113
Sequence 22, App1	959	5	1.0	24	3	US-09-012-126-11
Sequence 70, App1	960	5	1.0	24	3	US-09-276-202-10
Sequence 1, App1	961	5	1.0	24	4	PCT-US92-08094-12
Sequence 5, App1	962	5	1.0	24	4	PCT-US93-08364-9
Sequence 2, App1	963	5	1.0	24	4	PCT-US96-10905-5
Sequence 5, App1	964	5	1.0	25	1	US-08-233-146-5
Sequence 14, App1	965	5	1.0	25	1	US-08-463-470-5
Sequence 14, App1	966	5	1.0	25	2	US-08-792-555-25
Sequence 3, App1	967	5	1.0	25	5	5189345-24
Sequence 4, App1	968	5	1.0	26	1	US-07-942-245-425
Sequence 3, App1	969	5	1.0	26	1	US-07-942-245-457
Sequence 3, App1	970	5	1.0	26	1	US-08-560-727-5
Sequence 12, App1	971	5	1.0	26	1	US-08-634-600-26
Sequence 37, App1	972	5	1.0	26	2	US-08-870-846-9
Sequence 1, App1	973	5	1.0	26	2	US-08-818-255-9
Sequence 3, App1	974	5	1.0	26	3	US-08-818-255-9
Sequence 4, App1	975	5	1.0	26	3	US-08-256-747C-66
Sequence 5, App1	976	5	1.0	27	1	5262298-3
Sequence 3, App1	977	5	1.0	27	1	US-07-927-071-4
Sequence 3, App1	978	5	1.0	27	1	US-08-336-618-1

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-4

Query Match 46.6%; Score 243; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 8.3e-229;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGHYDNOLONGIRKRVKFELESSPNTQWELRAFAVAMHFSLTADRIDDILKVIYDSM 338
DB 87 ILKGHYDNOLONGIRKRVKFELESSPNTQWELRAFAVAMHFSLTADRIDDILKVIYDSM 146
QY 339 NHGDAARSKRELAELTALIKYVIAQEIINKHLSGGTINIHDSINIMDKNLGYTD 398
DB 147 NHGDAARSKRELAELTALIKYVIAQEIINKHLSGGTINIHDSINIMDKNLGYTD 206
QY 399 EEIFKASAEKIKLEKMPOTTIOYDGSEKIVSIKDFLGSEKRTGALGNKNSYKN 458
DB 207 EEIFKASAEKIKLEKMPOTTIOYDGSEKIVSIKDFLGSEKRTGALGNKNSYKN 266
QY 459 NELSHATTCSDKSRPLNDLVSOQTQLSDITSRNSAIEALNRFIOKDYVQRLDDT 518
DB 267 NELSHATTCSDKSRPLNDLVSOQTQLSDITSRNSAIEALNRFIOKDYVQRLDDT 326
QY 519 SGR 521
DB 327 SGR 329

RESULT 3
US-08-913-477-17
Sequence 17, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-17

Query Match 46.6%; Score 243; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.2e-228;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGHYDNOLONGIRKRVKFELESSPNTQWELRAFAVAMHFSLTADRIDDILKVIYDSM 338
DB 238 ILKGHYDNOLONGIRKRVKFELESSPNTQWELRAFAVAMHFSLTADRIDDILKVIYDSM 297
QY 339 NHGDAARSKRELAELTALIKYVIAQEIINKHLSGGTINIHDSINIMDKNLGYTD 398
DB 298 NHGDAARSKRELAELTALIKYVIAQEIINKHLSGGTINIHDSINIMDKNLGYTD 357
QY 399 EEIFKASAEKIKLEKMPOTTIOYDGSEKIVSIKDFLGSEKRTGALGNKNSYKN 458
DB 358 EEIFKASAEKIKLEKMPOTTIOYDGSEKIVSIKDFLGSEKRTGALGNKNSYKN 417
QY 459 NELSHATTCSDKSRPLNDLVSOQTQLSDITSRNSAIEALNRFIOKDYVQRLDDT 518
DB 418 NELSHATTCSDKSRPLNDLVSOQTQLSDITSRNSAIEALNRFIOKDYVQRLDDT 477
QY 519 SGR 521
DB 478 SGR 480

RESULT 4
US-08-913-477-23
Sequence 23, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-23

Query Match 46.6%; Score 243; DB 2; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.2e-228; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 ILKGHYNDQNLONGIKRYKEFLSSPNTQWELRAFMVHFSLTADRIDDDILKYIVDSM 338
DB 259 ILKGHYNDQNLONGIKRYKEFLSSPNTQWELRAFMVHFSLTADRIDDDILKYIVDSM 318
QY 339 NHHGDARSKLREBELAEIAELKIVSVIOAEIKHLSSTGCTINIHDKSINLMDKNLYGYTD 398
DB 319 NHHGDARSKLREBELAEIAELKIVSVIOAEIKHLSSTGCTINIHDKSINLMDKNLYGYTD 378
QY 399 EEIFASAEYKLEKMPOTTIOVDSEKKIVSIKDFLSENNKRTGALGTLKSYSYNKN 458
DB 379 EEIFASAEYKLEKMPOTTIOVDSEKKIVSIKDFLSENNKRTGALGTLKSYSYNKN 438
QY 459 NELSHFATTCSDKSRPLNDLVYSQKTTQSLDTSRFSNAIEALNRFIOKXDSVMQRLDDT 518
DB 439 NELSHFATTCSDKSRPLNDLVYSQKTTQSLDTSRFSNAIEALNRFIOKXDSVMQRLDDT 498
QY 519 SGK 521
DB 499 SGK 501

RESULT 5
US-08-913-477-21

; Sequence 21, Application US/08913477
; Patent No. 5985285

; GENERAL INFORMATION:

; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.

; APPLICANT: Leary, Sophie E.C.
; APPLICANT: Oyston, Petra C.F.

; APPLICANT: Bennett, Alice M.
; TITLE OF INVENTION: VACCINES FOR PLAGUE

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/913,477

FILING DATE: 15-SEP-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB96/00571

FILING DATE: 13-MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9505059.7

FILING DATE: 13-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9518946.0

FILING DATE: 15-SEP-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9524825.8

FILING DATE: 05-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Crawford, Arthur R.

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 124-599

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4100

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 170 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-913-477-21

Query Match 25.1%; Score 131; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 5.8e-120; Indels 0; Gaps 0;

Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MKKISSVIAIALFGIARANAADLASTATATVLEPARITLTYKEGAPITIMONGNIDT 83
DB 1 MKKISSVIAIALFGIARANAADLASTATATVLEPARITLTYKEGAPITIMONGNIDT 60
QY 84 ELVGTLLGKYKGTSTSVNFTDAAGDPYLTFTSODGNNHOFITKVGKDSRDPDIS 143
DB 61 ELVGTLLGKYKGTSTSVNFTDAAGDPYLTFTSODGNNHOFITKVGKDSRDPDIS 120
QY 144 PRVNGENLVGD 154
DB 121 PRVNGENLVGD 131

RESULT 6
US-08-913-477-11

; Sequence 11, Application US/08913477
; Patent No. 5985285

; GENERAL INFORMATION:

; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.

; APPLICANT: Leary, Sophie E.C.
; APPLICANT: Oyston, Petra C.F.

; APPLICANT: Bennett, Alice M.
; TITLE OF INVENTION: VACCINES FOR PLAGUE

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYTE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-477-11

Query Match 24.6%; Score 128; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 4.3e-117;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 TYREGAPITMDNGNITELLVGLITLGGYKGTSTSVNFTDAAGDPMTLTTSQDGN 125
DB 24 TYREGAPITMDNGNITELLVGLITLGGYKGTSTSVNFTDAAGDPMTLTTSQDGN 83

QY 126 HOFITKIVIGSDRFDISPRVNGENLVGDVYLATGSODFFVRISGSGKGLAAGXTDA 165
DB 84 HOFITKIVIGSDRFDISPRVNGENLVGDVYLATGSODFFVRISGSGKGLAAGXTDA 143

QY 186 VTVTVSNQ 193
DB 144 VTVTVSNQ 151

RESULT 7
US-08-780-496-8
Sequence 8, Application US/08780496
Patent No. 6046048
GENERAL INFORMATION:
APPLICANT: Avi Ashkenazi, Anan Chuntarapai, Kyung Jin Kim
TITLE OF INVENTION: Apo-2 Ligand
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California

COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/780,496
FILING DATE: 08-Jan-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maerschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P0978P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5416
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-08-780-496-8

Query Match 4.6%; Score 24; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNHHHHHHSSGHIDDDKHN 24
DB 1 MGNHHHHHHSSGHIDDDKHN 24

RESULT 8
US-08-651-818A-21
Sequence 21, Application US/08651818A
Patent No. 594889
GENERAL INFORMATION:
APPLICANT: de Boer, Plet A.J.
APPLICANT: Hale, Cynthia A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid

STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-21

Query Match 3.3%; Score 17; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.4e-10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 9
US-08-480-604A-24
Sequence 24, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
NUMBER OF SEQUENCES: 32
PREVENTION OF C. DIFFICILE DISEASE
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid

STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-604A-24

Query Match 3.3%; Score 17; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 10
US-08-405-496A-24
Sequence 24, Application US/08405496A
Patent No. 5913665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-405-496A-24

Query Match 3.3%; Score 17; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 11

US-08-651-818A-19
Sequence 19, Application US/08651818A
Patent No. 5948889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-19

Query Match 3.3%; Score 17; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 12

US-08-651-818A-23
Sequence 23, Application US/08651818A
Patent No. 5948889
GENERAL INFORMATION:
APPLICANT: de Boer, Piet A.J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America

ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,818A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02249
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-651-818A-23

Query Match 3.3%; Score 17; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 13

US-08-591-989-5
Sequence 5, Application US/08591989
Patent No. 5785721
GENERAL INFORMATION:
APPLICANT: Ross S. Rabin, Sumedha Jayasena
APPLICANT: and Larry Gold
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.40 MB
MEDIUM TYPE: storage
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,989
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX 49
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 323

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-591-989-5

Query Match 3.3%; Score 17; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGHNNHHHHSSGHI 17
Db 17 MGHNNHHHHSSGHI 33

RESULT 14
US-08-480-604A-26
Sequence 26, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHAY, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
NUMBER OF SEQUENCES: 32
PREVENTION OF C. DIFFICILE DISEASE
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-604A-26

Query Match 3.3%; Score 17; DB 1; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGHNNHHHHSSGHI 17
Db 1 MGHNNHHHHSSGHI 17

RESULT 15
US-08-405-496A-26
Sequence 26, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
NUMBER OF SEQUENCES: 30
NEUTROTOXIN
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-405-496A-26

Query Match 3.3%; Score 17; DB 2; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGHNNHHHHSSGHI 17
|||||

Wed Aug 23 11:46:14 2000

us-08-699-716a-2.oligo.ra1

Page 15

Db 1 MGHNNNNNNSSGRT 17

Search completed: August 22, 2000, 17:42:24
Job time: 1124 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:24:36 ; Search time 39.13 Seconds

(Without alignments)
823.970 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGHNNHHHHHHSSGHIDD.....RFIOKYDSVMQRLLDDTSGR 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 178050 seqs, 61884766 residues

Word size : 0

Total number of hits satisfying chosen parameters: 178050

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

PIR.64:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	46.6	326	2	T43594
2	191	36.7	326	2	B33601
3	170	32.6	170	2	S13008
4	82	15.7	326	2	B37314
5	5	3.3	168	2	I58315
6	11	2.1	302	2	A55641
7	11	2.1	303	2	A56837
8	11	2.1	303	2	B49122
9	11	2.1	303	2	A48130
10	11	2.1	354	2	S39406
11	11	2.1	355	2	S35345
12	11	2.1	156547	2	S41621
13	11	2.1	414	2	A40350
14	11	2.1	414	2	A48273
15	11	2.1	414	2	A56418
16	11	2.1	420	2	T39712
17	11	2.1	427	2	T42516
18	11	2.1	1180	2	S69205
19	11	1.9	132	2	I39193
20	10	1.9	133	2	B30242
21	10	1.9	222	2	I39192
22	10	1.9	236	2	S41512
23	10	1.9	264	2	S34121
24	10	1.9	298	2	S41469
25	10	1.9	309	2	T29293
26	10	1.9	321	2	T02987
27	10	1.9	330	2	I39194
28	10	1.9	331	2	A30242
29	10	1.9	335	2	G01448

30	10	1.9	351	1	K62GHL
31	10	1.9	375	2	A46390
32	10	1.9	387	2	A47446
33	10	1.9	390	1	A38565
34	10	1.9	410	2	I38502
35	10	1.9	411	2	I58156
36	10	1.9	430	2	S66671
37	10	1.9	451	2	A55909
38	10	1.9	469	2	I37451
39	10	1.9	475	2	A54743
40	10	1.9	480	2	T06672
41	10	1.9	580	2	T46024
42	10	1.9	623	1	S33167
43	10	1.9	625	2	S71930
44	10	1.9	628	2	JC2493
45	10	1.9	633	1	A26030
46	10	1.9	657	2	A29454
47	10	1.9	665	2	S02358
48	10	1.9	717	2	S38177
49	10	1.9	754	2	JC4898
50	10	1.9	786	2	A48456
51	10	1.9	950	2	S27473
52	10	1.9	1166	2	T13958
53	10	1.9	1196	2	T14108
54	10	1.9	1212	2	T13804
55	10	1.9	1249	2	T14270
56	10	1.9	1252	2	S36016
57	10	1.9	1273	2	T00338
58	10	1.9	1291	2	T13389
59	10	1.9	1293	2	T14259
60	10	1.9	1305	2	A40879
61	10	1.9	1312	1	B40879
62	10	1.9	1356	2	S51389
63	10	1.9	2212	2	A41098
64	9	1.7	143	1	B64421
65	9	1.7	192	2	T39367
66	9	1.7	255	2	S41511
67	9	1.7	270	2	A26480
68	9	1.7	307	2	A45581
69	9	1.7	311	2	A56235
70	9	1.7	331	2	S78452
71	9	1.7	332	2	JH0465
72	9	1.7	348	2	T04618
73	9	1.7	359	2	S14283
74	9	1.7	420	2	A49642
75	9	1.7	420	2	T59234
76	9	1.7	473	2	A54494
77	9	1.7	568	2	S15008
78	9	1.7	634	2	A54495
79	9	1.7	634	2	A28412
80	9	1.7	654	2	B71623
81	9	1.7	794	2	T39171
82	9	1.7	1116	2	S57382
83	9	1.7	1560	2	T00080
84	9	1.7	2254	2	T09053
85	9	1.7	2273	2	T46478
86	9	1.7	2273	2	T46478
87	9	1.7	2424	2	T46478
88	9	1.7	2424	2	T46478
89	9	1.7	60	2	C64698
90	9	1.5	83	2	T16435
91	8	1.5	102	2	T30119
92	8	1.5	115	2	H22583
93	8	1.5	140	2	A54523
94	8	1.5	143	2	E71524
95	8	1.5	208	2	E71524
96	8	1.5	216	2	E71524
97	8	1.5	250	2	T20257
98	8	1.5	292	2	T51171
99	8	1.5	323	2	T49529
100	8	1.5	323	2	T49529
101	8	1.5	351	2	T20270
102	8	1.5	386	2	S58755

histidine-rich gly
cAMP receptor subu
HNF-3/foxa head fa
polycarb (PC) prot
gene Brn-3b protei
Brn-3.2 - mouse
neuron-derived rec
transforming prote
HBF-G2 (HBF-2) pro
transcription fact
brain factor 1 pro
hypothetical prote
gene pointed prote
neuron-derived rec
serine/threonine-s
knob-associated hi
lamin A - Atrial
Sev7 protein homol
Down-syndrome-crit
occyat wall protei
URB51 protein - sm
SYNGAP-B1 protein
SH3-containing pro
shs protein - fruit
Ras-GTPase activat
occyat wall protei
hypothetical prote
hypothetical prote
ras GTPase-activat
phospholipase C (E
phospholipase C (E
ROM2 protein - yea
calcium channel pr
conserved hypothet
hypothetical prote
Brn-3a protein - m
knob protein - pla
Distal-less homeob
transcription acti
POU-domain protein
homeotic protein T
heat shock protein
transcription fact
transcription fact
octamer binding tr
knob-associated hi
gene disco protein
knob protein precu
histidine-rich pro
knob-associated hi
probable amine oxi
hypothetical prote
hypothetical prote
low voltage-activa
calcium channel BI
calcium channel BI
calcium channel BI
calcium channel BI
probable histidine
hypothetical prote
hypothetical prote
hypothetical prote
histidine-rich pro
hypothetical prote
ethand - mouse
hypothetical prote
transcription fact
transcription fact
homeotic protein o
hypothetical prote
ribosomal protein

103	8	1.5	401	2	S53405	probable membrane	176	7	1.3	321	2	T02936	carbonate dehydrat
104	8	1.5	409	2	S60988	hypothetical prote	177	7	1.3	328	2	S45998	hypothetical prote
105	8	1.5	471	2	T33997	hypothetical prote	178	7	1.3	328	2	G02469	homeotic protein D
106	8	1.5	476	2	S57963	methyl Cpg binding	179	7	1.3	329	1	S61884	carbonate dehydrat
107	8	1.5	495	1	S31223	transcription fact	180	7	1.3	330	1	S61883	carbonate dehydrat
108	8	1.5	499	2	S09880	hypothetical prote	181	7	1.3	330	1	S61882	carbonate dehydrat
109	8	1.5	508	2	S59870	fork head domain p	182	7	1.3	331	1	S48675	carbonate dehydrat
110	8	1.5	559	2	T26141	hypothetical prote	183	7	1.3	338	2	E75576	oxidoreductase - D
111	8	1.5	576	2	S69214	deformed epidermal	184	7	1.3	342	2	T06272	oxanessyl-pyrophosp
112	8	1.5	609	2	A49639	odd-paired - fruit	185	7	1.3	343	2	T29547	hypothetical prote
113	8	1.5	661	2	T46364	hypothetical prote	186	7	1.3	348	2	E64148	hypothetical prote
114	8	1.5	812	2	E75338	DNA gyrase, subunl	187	7	1.3	350	2	S76698	hypothetical prote
115	8	1.5	853	2	T46347	hypothetical prote	188	7	1.3	350	2	A30045	hypothetical prote
116	8	1.5	895	2	JC7089	zinc finger bindin	189	7	1.3	359	2	T21705	homeotic protein r
117	8	1.5	972	2	S35521	DNA topoisomerase	190	7	1.3	362	2	I40304	hypothetical prote
118	8	1.5	1133	2	T30302	P-type ATPase - Te	191	7	1.3	364	2	S11981	outer membrane lip
119	8	1.5	1145	2	T33606	hypothetical prote	192	7	1.3	374	2	T11662	variable major pro
120	8	1.5	1172	2	T00065	hypothetical prote	193	7	1.3	376	2	A49077	probable transcript
121	8	1.5	1203	2	T28895	hypothetical prote	194	7	1.3	380	2	S14188	transcription init
122	8	1.5	1441	2	T13717	CRAG protein - fru	195	7	1.3	385	1	OMRSP	carbonate dehydrat
123	8	1.5	1557	2	T13160	protein CNK - fru	196	7	1.3	391	2	A49645	seleoprotein P pr
124	8	1.5	185	2	T31611	hypothetical prote	197	7	1.3	392	1	A56691	transcription fact
125	8	1.5	1597	1	BVFSL	sol protein, large	198	7	1.3	394	2	A70979	phosphoribosylglic
126	8	1.5	1597	2	T08428	gene small optic l	199	7	1.3	398	2	G72333	hypothetical prote
127	8	1.5	1929	2	T21559	hypothetical prote	200	7	1.3	401	2	S09626	conserved hypothet
128	8	1.5	2416	2	T13825	adenomatous polypo	201	7	1.3	405	2	F81076	rib protein - Esc
129	8	1.5	5327	2	T13564	microtubule-associ	202	7	1.3	406	2	I41084	aspartokinase, alp
130	7	1.3	16	1	LEFCH	his operon leader	203	7	1.3	409	2	S40740	type I site-specific
131	7	1.3	77	2	T64836	hypothetical prote	204	7	1.3	410	2	S02167	cyclin A homolog -
132	7	1.3	83	2	C64893	gapc-2 protein, tr	205	7	1.3	415	2	T35834	probable transcript
133	7	1.3	86	2	T34975	hypothetical prote	206	7	1.3	418	2	D75283	hypothetical prote
134	7	1.3	91	1	BCROI7	S-100 protein beta	207	7	1.3	425	1	JH0710	probable phosphati
135	7	1.3	92	1	BCROI8	S-100 protein beta	208	7	1.3	434	2	I51436	hypothetical prote
136	7	1.3	92	2	A26557	S-100 protein beta	209	7	1.3	434	2	C71434	probable phosphati
137	7	1.3	104	1	G43258	S-100 protein beta	210	7	1.3	439	2	T46375	hypothetical prote
138	7	1.3	105	1	A23696	phosphotransferase	211	7	1.3	444	2	T09474	hypothetical prote
139	7	1.3	111	2	D53589	hypothetical prote	212	7	1.3	449	2	T10650	probable NADH dehy
140	7	1.3	111	2	S37150	hypothetical prote	213	7	1.3	456	2	UC4089	cationic amino aci
141	7	1.3	117	2	S42606	asir2 protein - tom	214	7	1.3	458	2	S57914	export protein - S
142	7	1.3	127	2	T44266	ribosomal protein	215	7	1.3	459	2	F75049	HMS1 protein - yea
143	7	1.3	136	2	T22959	hypothetical prote	216	7	1.3	466	2	T44746	methyl-Cpg-binding
144	7	1.3	137	2	T22959	hypothetical prote	217	7	1.3	470	2	F71686	serine/threonine-s
145	7	1.3	139	2	T33968	hypothetical prote	218	7	1.3	483	2	T35715	probable NADH dehy
146	7	1.3	142	2	T35319	probable regulator	219	7	1.3	490	2	S52830	export protein - S
147	7	1.3	147	2	D64331	hypothetical prote	220	7	1.3	492	2	A41907	HMS1 protein - yea
148	7	1.3	153	1	APBE21	endopeptidase (EC	221	7	1.3	513	1	S60303	serine/threonine-s
149	7	1.3	162	2	S22653	ribonuclease T2 (E	222	7	1.3	513	1	S60304	serine/threonine-s
150	7	1.3	162	2	S51002	ribonuclease precu	223	7	1.3	515	2	T38946	serine/threonine P
151	7	1.3	182	2	A53803	dihydrofolate redu	224	7	1.3	518	2	S44637	transducin homolo
152	7	1.3	182	2	H75176	probable dcm deam	225	7	1.3	519	2	A49367	hypothetical prote
153	7	1.3	182	2	E71037	hypothetical prote	226	7	1.3	529	2	T00677	hypothetical prote
154	7	1.3	187	2	T05071	hypothetical prote	227	7	1.3	532	2	S36554	probable transcript
155	7	1.3	193	2	H64174	hypothetical prote	228	7	1.3	532	2	T35119	hypothetical prote
156	7	1.3	209	2	T17825	hypothetical prote	229	7	1.3	535	1	A39194	probable aminotran
157	7	1.3	212	2	S41581	hypothetical prote	230	7	1.3	545	2	T11688	periplasmic dipept
158	7	1.3	236	2	S18783	exonoxin type A pr	231	7	1.3	550	2	T23760	phosphinositide-s
159	7	1.3	244	2	T07830	aluminum-induced p	232	7	1.3	564	2	T21983	hypothetical prote
160	7	1.3	252	2	S05629	cytochrome-c oxida	233	7	1.3	570	2	S07744	hypothetical prote
161	7	1.3	254	1	G64045	3-deoxy-manno-occu	234	7	1.3	579	2	JH0820	NADH dehydrogenase
162	7	1.3	255	2	T24883	hypothetical prote	235	7	1.3	587	2	T16617	160k gold antisen
163	7	1.3	259	2	A70359	hydrogenase expres	236	7	1.3	587	2	S12805	hypothetical prote
164	7	1.3	263	2	T16329	hypothetical prote	237	7	1.3	587	2	S41409	envelysin (EC 3.4.
165	7	1.3	264	2	T02886	carbonate dehydrat	238	7	1.3	601	2	T00119	probable transcript
166	7	1.3	273	2	T33424	hypothetical prote	239	7	1.3	604	2	S56027	hypothetical prote
167	7	1.3	281	2	E70388	conserved hypothet	240	7	1.3	615	2	S54468	hypothetical prote
168	7	1.3	284	2	T35723	cobalt transport s	241	7	1.3	617	2	S54468	hypothetical prote
169	7	1.3	290	2	T21868	hypothetical prote	242	7	1.3	620	2	S52494	hypothetical prote
170	7	1.3	294	2	E69759	hypothetical prote	243	7	1.3	622	2	S71342	5-methyltetrahydro
171	7	1.3	306	2	A45335	cytochrome-c oxida	244	7	1.3	627	2	T37293	protein kinase hom
172	7	1.3	306	2	T09067	extensin-like prot	245	7	1.3	642	2	T39141	choleline o-acetyltr
173	7	1.3	314	2	T29363	hypothetical prote	246	7	1.3	682	2	A44493	hypothetical prote
174	7	1.3	314	2	E64580	cathepsin-like prot	247	7	1.3	688	2	T23108	serum-inducible ki
175	7	1.3	314	2	E71931	hypothetical prote	248	7	1.3	701	2	T20892	hypothetical prote

249	7	1.3	732	2	547073	finger protein HZF
250	7	1.3	733	2	H69411	cell division gont
251	7	1.3	736	2	H75460	conserved hypotnet
252	7	1.3	746	2	G65100	formate C-acetyltr
253	7	1.3	766	2	G71437	probable resistant
254	7	1.3	778	2	T44761	probable preprotel
255	7	1.3	785	2	S57702	hypothetical prote
256	7	1.3	794	2	F72202	hypothetical prote
257	7	1.3	807	2	T42924	glycoprotein B - a
258	7	1.3	808	2	F70720	probable transloca
259	7	1.3	808	2	J02205	Ula7h protein - Ma
260	7	1.3	845	2	T17291	hypothetical prote
261	7	1.3	886	1	UC5085	replication licens
262	7	1.3	892	1	S42228	replication licens
263	7	1.3	892	2	S76350	hypothetical prote
264	7	1.3	901	2	JC7111	tetratricopeptide
265	7	1.3	904	1	T10067	replication licens
266	7	1.3	906	1	RNBPK1	DNA-directed RNA p
267	7	1.3	906	2	A71438	probable resistant
268	7	1.3	906	2	A43817	transforming prote
269	7	1.3	944	2	C70839	probable mmp13 pro
270	7	1.3	946	2	S48255	probable membrane
271	7	1.3	946	2	S48433	inositol-1,4,5-tri
272	7	1.3	1001	2	T28897	hypothetical prote
273	7	1.3	1133	2	T13384	hypothetical prote
274	7	1.3	1133	2	T13246	hypothetical prote
275	7	1.3	1150	2	T13824	letthal(1)1b1 prote
276	7	1.3	1176	2	T18042	LK6 protein kinase
277	7	1.3	1289	2	S67200	ice nucleation pro
278	7	1.3	1314	2	G02870	hypothetical prote
279	7	1.3	1356	2	T18521	KIAA0197 protein -
280	7	1.3	1361	2	T29435	beta transducin-11
281	7	1.3	1465	2	T23056	hypothetical prote
282	7	1.3	1531	2	H71468	hypothetical prote
283	7	1.3	1691	1	D54689	probable outer mem
284	7	1.3	1705	2	S51672	protein-tyrosine-p
285	7	1.3	1706	1	OBRC	adenylate cyclase
286	7	1.3	1894	2	C54689	cyclolysin - Borde
287	7	1.3	1912	2	A56178	protein-tyrosine-p
288	7	1.3	1979	2	C71622	protein-tyrosine-p
289	7	1.3	2117	2	T36180	hypothetical prote
290	7	1.3	2226	2	B47447	CD4 peptide synthe
291	7	1.3	2467	2	D71437	calcium channel pr
292	7	1.3	3104	2	S20473	probable resistant
293	7	1.3	3345	2	T13423	fatly-acid synthas
294	7	1.3	3413	3	T17467	hypothetical prote
295	7	1.3	3413	3	A61328	rifamycin polyketi
296	6	1.2	30	2	S15678	trypsin (EC 3.4.21
297	6	1.2	31	2	S20491	acetylcholinestera
298	6	1.2	31	2	T07276	hypothetical prote
299	6	1.2	34	2	A39965	photosystem II pro
300	6	1.2	35	2	A05057	macrocclin-O-methyl
301	6	1.2	53	2	A56512	hypothetical prote
302	6	1.2	54	2	A31440	nuclear pore prote
303	6	1.2	54	2	C31436	ovomucoid, third d
304	6	1.2	57	2	A61587	ovomucoid, third d
305	6	1.2	54	2	H64698	histidine and glut
306	6	1.2	62	2	B69788	hypothetical prote
307	6	1.2	64	2	S57787	hypothetical prote
308	6	1.2	65	2	S01293	acetylcholinestera
309	6	1.2	65	2	S48113	transcription fact
310	6	1.2	65	2	T45078	hypothetical prote
311	6	1.2	65	2	E75022	ssu ribosomal prot
312	6	1.2	66	2	T35489	hypothetical prote
313	6	1.2	66	2	T00821	hypothetical prote
314	6	1.2	70	2	E70985	hypothetical prote
315	6	1.2	71	2	A39566	hypothetical prote
316	6	1.2	76	2	T06117	H+-transporting At
317	6	1.2	77	2	D71821	hypothetical prote
318	6	1.2	77	2	C75495	probable histidine
319	6	1.2	78	2	T21426	hypothetical prote
320	6	1.2	79	2	B72223	hypothetical prote
321	6	1.2	82	2	S22306	hypothetical prote
322	6	1.2	83	2	T48283	gene CC2 protein -
323	6	1.2	83	2	C75083	hypothetical prote
324	6	1.2	85	2	S24453	hypothetical prote
325	6	1.2	87	2	T25505	hypothetical prote
326	6	1.2	89	1	TKMY	tuberculin-active
327	6	1.2	89	2	C72286	transcription regu
328	6	1.2	89	2	T14959	hypothetical prote
329	6	1.2	89	2	S77075	hypothetical prote
330	6	1.2	91	2	T41989	hypothetical prote
331	6	1.2	92	1	Z1BPC2	regulatory protein
332	6	1.2	92	2	B71637	hypothetical prote
333	6	1.2	92	2	T34146	hypothetical prote
334	6	1.2	92	2	T20721	hypothetical prote
335	6	1.2	93	2	T31266	ferredoxin [2Fe-2S
336	6	1.2	95	2	S11788	nolp protein - Rhl
337	6	1.2	99	1	MGRB2	beta-2-microglobul
338	6	1.2	100	2	B71567	hypothetical prote
339	6	1.2	100	2	G81778	conserved hypotnet
340	6	1.2	101	2	B44971	hypothetical prote
341	6	1.2	103	1	WQSA3L	phosphotransferase
342	6	1.2	106	2	A03860	hypothetical prote
343	6	1.2	108	2	S34940	fimbrial protein S
344	6	1.2	108	2	S34942	fimbrial protein S
345	6	1.2	108	2	A45649	microfilament surf
346	6	1.2	109	2	G71077	hypothetical prote
347	6	1.2	111	2	S07367	hypothetical prote
348	6	1.2	112	1	ASLJGH	vpu protein - huma
349	6	1.2	112	2	S12155	hypothetical 13.2K
350	6	1.2	112	2	S03221	conserved hypotnet
351	6	1.2	112	2	H69537	photosystem I 8.4K
352	6	1.2	113	2	S06684	B27R protein - vac
353	6	1.2	113	2	H42528	hypothetical prote
354	6	1.2	115	2	F75140	hypothetical prote
355	6	1.2	117	2	T07281	MDH dehydrogenase
356	6	1.2	118	2	T20862	hypothetical prote
357	6	1.2	118	2	H81443	diacylglycerol kin
358	6	1.2	120	2	E71160	probable chemotaxi
359	6	1.2	120	2	F75002	chemotaxis respons
360	6	1.2	121	1	LAKGAW	alpha-lactalbumin
361	6	1.2	122	1	KRECDG	probable diacygly
362	6	1.2	122	2	T01558	auxin-induced prot
363	6	1.2	123	1	S26183	general stress pro
364	6	1.2	125	2	S24831	hypothetical prote
365	6	1.2	127	2	A72388	hypothetical prote
366	6	1.2	127	2	E65098	hypothetical 14.2K
367	6	1.2	130	2	S14983	extensin class I (
368	6	1.2	131	1	C70430	conserved hypotnet
369	6	1.2	132	1	MEIV2J	matrix protein M2
370	6	1.2	133	2	A64661	biopolymer transpo
371	6	1.2	133	2	A71855	biopolymer transpo
372	6	1.2	134	2	A26860	acyl carrier prote
373	6	1.2	134	2	S20499	acyl carrier prote
374	6	1.2	134	2	B71392	hypothetical 2b pr
375	6	1.2	135	1	WABP77	gene 4.7 protein -
376	6	1.2	137	2	S38478	hypothetical prote
377	6	1.2	137	2	F75043	hypothetical prote
378	6	1.2	139	3	T41428	WD repeat protein
379	6	1.2	140	1	WMVZM3	HM3 protein - shie
380	6	1.2	140	2	A60394	alpha-lactalbumin
381	6	1.2	140	2	T06554	probable profilin
382	6	1.2	140	2	T19083	hypothetical prote
383	6	1.2	141	2	T06553	probable profilin
384	6	1.2	142	1	R3178	ribosomal protein
385	6	1.2	142	1	F70234	hypothetical prote
386	6	1.2	143	2	T14177	ribosomal protein
387	6	1.2	143	2	T16332	hypothetical prote
388	6	1.2	144	1	TPBPPI	tail fiber protein
389	6	1.2	144	2	T11756	transcription regu
390	6	1.2	144	2	T42287	hypothetical prote
391	6	1.2	144	2	A69971	conserved hypotnet
392	6	1.2	144	2	T25028	hypothetical prote
393	6	1.2	146	2	F64449	hypothetical prote
394	6	1.2	146	2	T07125	plastid-lipid-asso

395	6	1.2	147	2	T46623	rearranged T-cell	468	6	1.2	190	2	G69339	conserved hypother
396	6	1.2	147	2	G72384	conserved hypother	469	6	1.2	191	2	S43178	gene s122 protein
397	6	1.2	148	2	T16177	hypothetical prote	470	6	1.2	191	2	T32278	hypothetical prote
398	6	1.2	148	2	JN0568	single-stranded m	471	6	1.2	192	2	JN0577	hypothetical prote
399	6	1.2	148	2	C75126	hypothetical prote	472	6	1.2	192	2	F70222	p-aminobenzoic aci
400	6	1.2	149	2	H72344	conserved hypother	473	6	1.2	193	2	T36959	conserved hypother
401	6	1.2	149	2	F70599	hypothetical prote	474	6	1.2	193	2	C57256	hypothetical prote
402	6	1.2	150	2	T23641	hypothetical prote	475	6	1.2	193	2	E75256	hypothetical prote
403	6	1.2	150	2	T12547	hypothetical prote	476	6	1.2	194	2	D69486	Uag-related protei
404	6	1.2	151	2	JN0569	single-stranded m	477	6	1.2	194	2	C44805	probable ribosomal
405	6	1.2	151	2	C41700	cl protein - rabbl	478	6	1.2	195	2	I46272	eggshell protein -
406	6	1.2	151	2	C75157	protein tyrosine p	479	6	1.2	195	2	F64842	trophoblast interf
407	6	1.2	153	2	G65108	transcription elon	480	6	1.2	197	1	A65081	hypothetical prote
408	6	1.2	154	2	H69320	hypothetical prote	481	6	1.2	197	2	S25825	hypothetical prote
409	6	1.2	154	2	T13439	hypothetical prote	482	6	1.2	197	2	H71003	polysaccharonase
410	6	1.2	157	2	H69203	conserved hypother	483	6	1.2	198	2	S73275	hypothetical prote
411	6	1.2	158	2	F70955	hypothetical prote	484	6	1.2	198	2	B71214	hypothetical prote
412	6	1.2	158	2	T05097	hypothetical prote	485	6	1.2	198	2	S24335	hypothetical prote
413	6	1.2	158	2	D75069	hypothetical prote	486	6	1.2	198	2	F55578	H+-transporting Ar
414	6	1.2	158	2	B71096	hypothetical prote	487	6	1.2	199	2	C75363	hypothetical prote
415	6	1.2	159	2	S72326	glucan 1,3-beta-gl	488	6	1.2	199	2	T05283	hydrolase family p
416	6	1.2	159	2	S33585	lipoprotein smpa p	489	6	1.2	199	2	S59741	hypothetical prote
417	6	1.2	161	2	B81357	probable lipoprote	490	6	1.2	200	2	S75359	PAC10 protein - ye
418	6	1.2	161	2	E43719	uree protein - Pro	491	6	1.2	200	2	T36125	hypothetical prote
419	6	1.2	162	2	A27398	allophycocyanin al	492	6	1.2	201	2	T38713	H+-transporting Ar
420	6	1.2	162	2	T45343	hypothetical prote	493	6	1.2	202	2	S72224	hypothetical prote
421	6	1.2	162	2	H70927	probable rpsp prot	494	6	1.2	202	2	T18495	hypothetical prote
422	6	1.2	162	2	T33127	hypothetical prote	495	6	1.2	203	2	S67607	hypothetical prote
423	6	1.2	163	2	I40858	superoxide dismuta	496	6	1.2	203	2	T33126	probable membrane
424	6	1.2	163	2	S28136	gas vesicle protei	497	6	1.2	204	2	B75539	hypothetical prote
425	6	1.2	164	1	ZPCL	lipoprotein signal	498	6	1.2	204	2	T34944	probable phosphogl
426	6	1.2	164	1	T24726	hypothetical prote	499	6	1.2	204	2	S67295	probable lipoprote
427	6	1.2	167	2	B70458	lipoprotein signal	500	6	1.2	204	2	D75154	probable membrane
428	6	1.2	168	2	I40346	ompl6 protein - Br	501	6	1.2	205	2	E69327	hypothetical prote
429	6	1.2	168	2	C71450	hypothetical prote	502	6	1.2	205	2	E75310	ribonuclease HII (
430	6	1.2	168	2	T25336	hypothetical prote	503	6	1.2	206	2	T33064	hypothetical prote
431	6	1.2	168	2	A70479	hypothetical prote	504	6	1.2	206	2	T25384	hypothetical prote
432	6	1.2	169	1	S23478	probable benzozate	505	6	1.2	208	2	T24446	hypothetical prote
433	6	1.2	170	2	B37152	lipoprotein signal	506	6	1.2	208	2	D71258	probable gipg prot
434	6	1.2	172	2	H65026	hypothetical prote	507	6	1.2	209	2	H81149	conserved hypother
435	6	1.2	172	2	T36107	probable serine/ar	508	6	1.2	210	2	S25657	T-cell surface gly
436	6	1.2	172	2	C72007	low calcium respon	509	6	1.2	210	2	S64376	hypothetical prote
437	6	1.2	172	2	G81533	type III secretion	510	6	1.2	210	2	JC5512	hypothetical prote
438	6	1.2	173	1	YOECP	fibritial protein p	511	6	1.2	211	2	B71809	SNARE protein 23 -
439	6	1.2	173	2	I41201	E-minor fibritial p	512	6	1.2	211	2	E64710	hypothetical prote
440	6	1.2	173	2	B27743	pape fibritial prot	513	6	1.2	211	2	C59091	hypothetical prote
441	6	1.2	173	2	S25210	prse protei - Esc	514	6	1.2	211	2	T00501	hypothetical prote
442	6	1.2	173	2	C81549	conserved hypother	515	6	1.2	213	2	E64690	RER1C protein homo
443	6	1.2	174	2	JC1294	pape fibritial prot	516	6	1.2	213	2	C68641	response regulator
444	6	1.2	175	2	C27743	pape fibritial prot	517	6	1.2	213	2	S75247	ATP phosphoribosyl
445	6	1.2	175	2	C69498	conserved hypother	518	6	1.2	214	2	I49758	hypothetical prote
446	6	1.2	175	2	T46136	remorin-like prote	519	6	1.2	214	2	D64163	hypoxanthine phosph
447	6	1.2	175	2	I53285	glucocorticoid-reg	520	6	1.2	214	2	T47110	hypothetical prote
448	6	1.2	175	2	A53523	endocrine secretor	521	6	1.2	215	2	B75086	3-oxoadipate CoA-t
449	6	1.2	175	2	JC5912	lysyl protein - Lac	522	6	1.2	216	2	T26054	hypothetical prote
450	6	1.2	175	2	T13323	protein R175 - Lac	523	6	1.2	218	1	RTMSG	hypoxanthine phosph
451	6	1.2	177	2	D71869	hypothetical prote	524	6	1.2	218	1	RTHYG	hypoxanthine phosph
452	6	1.2	178	2	H81207	conserved hypother	525	6	1.2	218	2	S43043	hypoxanthine (guan
453	6	1.2	179	1	R5B55	ribosomal protein	526	6	1.2	218	2	S21474	hypoxanthine phosph
454	6	1.2	179	2	S29313	inorganic pyrophos	527	6	1.2	218	2	T51842	hypoxanthine phosph
455	6	1.2	179	2	B75636	transcription regu	528	6	1.2	218	2	S16140	hypoxanthine phosph
456	6	1.2	180	2	F75073	pyrazinamidase/nic	529	6	1.2	218	2	T26826	hypoxanthine phosph
457	6	1.2	181	2	H69502	hypothetical prote	530	6	1.2	219	1	C64625	probable thiamin-p
458	6	1.2	182	2	G70698	hypothetical prote	531	6	1.2	219	2	T32443	hypothetical prote
459	6	1.2	182	2	A02946	keratin, 59K type	532	6	1.2	220	2	B53312	hypothetical prote
460	6	1.2	182	2	A02947	keratin, 60K type	533	6	1.2	220	2	A44805	deoxyribose-phosph
461	6	1.2	183	2	T27236	hypothetical prote	534	6	1.2	220	2	T44702	eggshell protein -
462	6	1.2	184	2	S14688	hypothetical prote	535	6	1.2	221	2	G73322	hypothetical prote
463	6	1.2	184	2	T06673	hypothetical prote	536	6	1.2	221	2	A70045	glutaredoxin-relat
464	6	1.2	184	2	T24251	hypothetical prote	537	6	1.2	221	2	T07079	pectate lyase homo
465	6	1.2	184	2	PC4032	nitrite reductase	538	6	1.2	222	2	T29457	leucine-rich repea
466	6	1.2	186	2	G72373	lipoprotein signal	539	6	1.2	222	2	S34563	gene M1 protein -
467	6	1.2	188	2	S67051	hypothetical prote	540	6	1.2	222	2	B49599	polymerase-associa

541	6	1.2	223	1	QOBE26	614	6	1.2	256	2	D69845	thiamin biosynthes
542	6	1.2	223	2	T07098	615	6	1.2	256	2	A69129	conserved hypotet
543	6	1.2	224	1	KICHT	616	6	1.2	257	2	A28664	enterotoxin A prec
544	6	1.2	224	2	B70327	617	6	1.2	257	2	D71544	hypothetical prote
545	6	1.2	224	2	T10660	618	6	1.2	258	1	F69374	conserved hypotet
546	6	1.2	226	1	G64387	619	6	1.2	258	2	T01481	carbonate dehydrat
547	6	1.2	226	2	C71307	620	6	1.2	258	2	D70758	hypothetical prote
548	6	1.2	226	2	A70565	621	6	1.2	258	2	B72697	hypothetical prote
549	6	1.2	226	2	G69129	622	6	1.2	260	2	B64563	hypothetical prote
550	6	1.2	227	2	T11080	623	6	1.2	261	2	S10321	3ik protein - frog
551	6	1.2	227	2	T46686	624	6	1.2	261	2	F81376	probable Integral
552	6	1.2	230	2	T44953	625	6	1.2	262	2	F72333	hypothetical prote
553	6	1.2	230	2	T18757	626	6	1.2	263	2	G71494	probable php hydrol
554	6	1.2	231	1	F1SP3	627	6	1.2	263	2	B81653	mtc protein, prob
555	6	1.2	232	2	S77017	628	6	1.2	264	2	G69632	flagellar hook pro
556	6	1.2	232	2	S75919	629	6	1.2	264	2	S66439	hypothetical 29.7k
557	6	1.2	234	2	S22360	630	6	1.2	265	1	B65132	hypothetical trans
558	6	1.2	235	2	S60158	631	6	1.2	265	2	F70907	hypothetical prote
559	6	1.2	235	2	H72093	632	6	1.2	266	2	T136287	probable Integral
560	6	1.2	235	2	D81576	633	6	1.2	267	2	S76499	hypothetical prote
561	6	1.2	236	2	D69423	634	6	1.2	268	2	G69623	flagellar hook-bas
562	6	1.2	237	2	B75214	635	6	1.2	268	2	B39439	CAMP response elem
563	6	1.2	237	2	T35855	636	6	1.2	268	2	C37280	enhancer-binding p
564	6	1.2	237	2	C40784	637	6	1.2	268	2	A56446	C/EBP-related prot
565	6	1.2	239	1	T35279	638	6	1.2	269	2	A56446	ig heavy chain V r
566	6	1.2	239	1	B69406	639	6	1.2	269	2	A40225	transcription acti
567	6	1.2	239	2	T20255	640	6	1.2	269	2	A47008	transcription acti
568	6	1.2	239	2	S59050	641	6	1.2	269	2	S51814	hypothetical prote
569	6	1.2	239	2	B72573	642	6	1.2	270	2	S75051	lactose transport
570	6	1.2	240	2	T45727	643	6	1.2	270	2	T03635	plastid-lipid-asso
571	6	1.2	240	2	B71615	644	6	1.2	270	2	JE0167	nitric-oxide reduc
572	6	1.2	241	2	F71478	645	6	1.2	271	2	S25006	calretinin - rat
573	6	1.2	241	2	E81741	646	6	1.2	271	2	A60253	transcription fact
574	6	1.2	242	2	D81438	647	6	1.2	271	2	JC6553	hypothetical prote
575	6	1.2	242	2	E71621	648	6	1.2	271	2	D70591	granaticin polyket
576	6	1.2	242	2	T21226	649	6	1.2	271	2	T01767	membrane protein 1
577	6	1.2	243	1	JJAG32	650	6	1.2	272	1	S05397	hypothetical prote
578	6	1.2	243	1	S06738	651	6	1.2	273	1	MMAGCG	type II site-speci
579	6	1.2	243	2	S03122	652	6	1.2	273	2	E64316	hypothetical prote
580	6	1.2	243	2	D71450	653	6	1.2	273	2	JC1166	type II site-speci
581	6	1.2	243	2	T12734	654	6	1.2	274	1	JU0957	myb-related protei
582	6	1.2	244	2	T31838	655	6	1.2	274	2	S22308	type II site-speci
583	6	1.2	245	2	A55456	656	6	1.2	274	2	A64978	hypothetical prote
584	6	1.2	245	2	T24565	657	6	1.2	275	2	T02334	hypothetical prote
585	6	1.2	246	1	TRDGC	658	6	1.2	277	2	B69479	conserved accessory p
586	6	1.2	246	2	I50209	659	6	1.2	278	2	F72064	solute binding pro
587	6	1.2	247	2	S53863	660	6	1.2	278	2	T33129	hypothetical prote
588	6	1.2	247	2	A43846	661	6	1.2	278	2	B81601	ABC transporter, p
589	6	1.2	247	2	T29733	662	6	1.2	279	2	G72637	hypothetical prote
590	6	1.2	248	2	T40343	663	6	1.2	280	2	H72389	transcription regu
591	6	1.2	248	2	S64486	664	6	1.2	280	2	S42479	hypothetical prote
592	6	1.2	248	2	T21786	665	6	1.2	281	2	S44230	stir protein - Str
593	6	1.2	249	2	G64876	666	6	1.2	282	2	C64435	shikimate 5-dehydr
594	6	1.2	249	2	T21785	667	6	1.2	283	2	S52928	hypothetical prote
595	6	1.2	249	2	G75189	668	6	1.2	283	2	T31122	probable transcrip
596	6	1.2	250	2	T20256	669	6	1.2	284	2	G72107	oligopeptide trans
597	6	1.2	250	2	T20258	670	6	1.2	284	2	F81563	peptide ABC transp
598	6	1.2	250	2	F72057	671	6	1.2	285	2	S78841	hypothetical prote
599	6	1.2	250	2	C64304	672	6	1.2	285	2	C72516	hypothetical prote
600	6	1.2	250	2	E81609	673	6	1.2	285	2	T21962	hypothetical prote
601	6	1.2	252	1	S59047	674	6	1.2	286	2	C64538	cytochrome-c oxida
602	6	1.2	252	1	REPCDO	675	6	1.2	286	2	T10129	thymidylate synth
603	6	1.2	252	2	T40604	676	6	1.2	287	2	F70935	probable ribosomal
604	6	1.2	252	2	T24939	677	6	1.2	287	2	A47038	prolyl aminopeptid
605	6	1.2	252	2	A32489	678	6	1.2	288	2	B41768	orotidine-5'-phosp
606	6	1.2	253	2	B71346	679	6	1.2	288	2	T40472	probable splindle p
607	6	1.2	253	2	F64741	680	6	1.2	289	2	C70400	ferredoxin oxidore
608	6	1.2	253	2	T15385	681	6	1.2	289	2	A69161	hypothetical prote
609	6	1.2	253	2	B70483	682	6	1.2	289	2	S16556	hypothetical prote
610	6	1.2	255	1	S04899	683	6	1.2	289	2	S13800	centrosomin A - mo
611	6	1.2	255	1	E72360	684	6	1.2	290	2	S55276	N-acetyltransferas
612	6	1.2	255	2	A75502	685	6	1.2	291	1	EWMTG	gamma-glutadin B pr
613	6	1.2	256	2	S37893	686	6	1.2	292	2	C75024	gamma-glutadinase, gamma

687	1.2	293	2	B32010	760	1.2	325	2	S75747	hypothetical prote
688	1.2	293	2	F73564	761	1.2	325	2	G75553	conserved hypothet
689	1.2	293	2	G69954	762	1.2	325	2	A69930	conserved hypothet
690	1.2	293	2	T31840	763	1.2	326	2	C70645	hypothetical prote
691	1.2	294	2	S73949	764	1.2	326	2	T09548	phosphoprotein pho
692	1.2	295	2	S46708	765	1.2	326	2	S57544	probable membrane
693	1.2	296	1	OPNE7	766	1.2	326	2	E64887	hypothetical prote
694	1.2	296	2	T05192	767	1.2	326	2	T04055	hypothetical prote
695	1.2	296	2	T32222	768	1.2	326	2	T07825	fibritillin homolog
696	1.2	296	2	F72745	769	1.2	327	2	T45143	hypothetical prote
697	1.2	297	2	S23737	770	1.2	328	2	S06780	hypothetical prote
698	1.2	297	2	T00180	771	1.2	328	2	S15916	hypothetical prote
699	1.2	298	2	E70932	772	1.2	328	2	S38138	transcription init
700	1.2	298	2	E71200	773	1.2	329	2	T30513	hypothetical prote
701	1.2	298	2	G64465	774	1.2	329	2	T45972	hypothetical prote
702	1.2	299	2	F75003	775	1.2	330	2	F71146	hypothetical prote
703	1.2	299	2	D65187	776	1.2	332	1	T50074	biotin synthetase
704	1.2	299	2	T25435	777	1.2	332	2	A22366	glyceraldenhyde-3-p
705	1.2	301	1	A54687	778	1.2	332	2	B22366	glyceraldenhyde-3-p
706	1.2	301	1	A48880	779	1.2	332	2	S24630	glyceraldenhyde-3-p
707	1.2	301	1	I56543	780	1.2	332	2	S64375	probable tyrosine
708	1.2	301	2	S28708	781	1.2	332	2	C72310	conserved hypothet
709	1.2	301	2	JC5348	782	1.2	332	2	E81665	protease IV, pro
710	1.2	301	2	H71011	783	1.2	333	2	A39098	allergen Poa p IX
711	1.2	302	2	JA0153	784	1.2	333	2	A72371	oligopeptide ABC t
712	1.2	302	2	T15706	785	1.2	334	2	T41705	probable 2-hydroxy
713	1.2	303	1	GEPO	786	1.2	334	2	T42743	hypothetical prote
714	1.2	303	1	B69160	787	1.2	334	2	T39602	glyceraldenhyde-3-p
715	1.2	303	2	S61723	788	1.2	334	2	H42522	AS1r protein - vac
716	1.2	304	2	A30106	789	1.2	334	2	Q41789	Salp14r protein -
717	1.2	304	2	D70370	790	1.2	334	2	S46845	USF protein - vari
718	1.2	305	2	D69199	791	1.2	334	2	T28596	27r protein - vari
719	1.2	305	2	A56554	792	1.2	334	2	F64674	aliphatic amidase
720	1.2	305	2	F70533	793	1.2	334	2	C71842	hypothetical prote
721	1.2	305	2	I57039	794	1.2	334	2	B72171	uroporhyrinogen d
722	1.2	305	2	D70432	795	1.2	336	2	T09186	UDP-glucose 4-epim
723	1.2	306	2	S25824	796	1.2	336	2	G69270	phenylalanine--trn
724	1.2	306	2	E64184	797	1.2	336	2	C64410	probable transposa
725	1.2	306	2	G64820	798	1.2	338	1	IEEC5D	hypothetical prote
726	1.2	306	2	JC1120	799	1.2	338	2	G70329	probable transposa
727	1.2	307	2	B72031	800	1.2	338	2	S75089	uroporhyrinogen d
728	1.2	307	2	S38152	801	1.2	338	2	D70382	UDP-glucose 4-epim
729	1.2	307	2	T18483	802	1.2	338	2	A64964	phenylalanine--trn
730	1.2	307	2	C81509	803	1.2	339	2	A56274	probable transposa
731	1.2	308	2	E70392	804	1.2	339	2	E69717	sulfur-regulated 3
732	1.2	308	2	E81288	805	1.2	340	2	E75436	spore coat polysac
733	1.2	308	2	G70696	806	1.2	340	2	T35394	conserved hypothet
734	1.2	308	2	T20043	807	1.2	340	2	S65000	probable glycerol
735	1.2	309	2	T09166	808	1.2	341	2	T38439	probable membrane
736	1.2	310	2	S36581	809	1.2	341	2	F72501	electron transfer
737	1.2	310	2	T37437	810	1.2	342	1	B71051	probable eukaryoti
738	1.2	311	2	S52498	811	1.2	342	2	C70401	conserved hypothet
739	1.2	312	2	H64560	812	1.2	344	2	A43592	phospho-N-acetylm
740	1.2	312	2	T31834	813	1.2	344	2	T15457	outer membrane pro
741	1.2	312	2	A56911	814	1.2	345	1	S75490	hypothetical prote
742	1.2	314	2	H70840	815	1.2	345	1	MNKMT	probable hydro-ly
743	1.2	314	2	E81064	816	1.2	345	1	T12360	nonstructural prot
744	1.2	315	2	S60918	817	1.2	345	2	S73729	NADH dehydrogenase
745	1.2	315	2	T10252	818	1.2	346	2	D64680	MG307 homolog H08
746	1.2	315	2	T09165	819	1.2	347	2	S53348	conserved hypothet
747	1.2	316	2	A38265	820	1.2	347	2	S64936	nuclear fusion pro
748	1.2	317	2	G69285	821	1.2	348	1	PABY	fructose-bisphosph
749	1.2	317	2	G70313	822	1.2	348	2	B75097	hypothetical prote
750	1.2	318	2	E71240	823	1.2	348	2	F71372	probable protoporp
751	1.2	320	2	A43669	824	1.2	349	2	T42416	probable electron
752	1.2	320	2	T21452	825	1.2	349	2	A64090	asparaginase (EC 3
753	1.2	321	2	B72384	826	1.2	349	2	I69731	hypothetical prote
754	1.2	322	2	S24753	827	1.2	349	2	T46362	probable tyrosine
755	1.2	322	2	S75280	828	1.2	349	2	H64456	hypothetical prote
756	1.2	322	2	E81279	829	1.2	350	2	S61414	DNA-binding protei
757	1.2	323	2	E65109	830	1.2	350	2	T06526	UDPglucose 4-epim
758	1.2	324	2	F69287	831	1.2	350	2	D70805	hypothetical prote
759	1.2	324	2	S39502	832	1.2	351	2	E71935	UDPurotate dehy

833	6	1.2	351	2	E75438	lytrophanyl-tRNA
834	6	1.2	351	2	B72315	conserved hypotnet
835	6	1.2	352	2	JS0023	glycerol-3-phospha
836	6	1.2	352	2	T20729	hypothetical prote
837	6	1.2	352	2	D69133	coenzyme F420-redu
838	6	1.2	352	2	T18794	hypothetical prote
839	6	1.2	353	2	S11790	glycerol-3-phospha
840	6	1.2	353	2	C31751	protein kinase (EC
841	6	1.2	353	2	T35921	probable oxidoredu
842	6	1.2	354	2	T36559	recombinant epoxide h
843	6	1.2	355	1	ROEBPM	recombination prot
844	6	1.2	355	1	S30315	DNA-methyltransfer
845	6	1.2	355	2	S41686	geranylgeranyltran
846	6	1.2	355	2	S11152	oligopeptide trans
847	6	1.2	355	2	C39725	hypothetical prote
848	6	1.2	356	2	T41764	ACKNP orf18 - Bom
849	6	1.2	358	2	T46532	protein co-factor
850	6	1.2	359	1	DEUT1C	glycerol-3-ph
851	6	1.2	359	1	DEUT1B	glycerol-3-ph
852	6	1.2	359	2	T32287	hypothetical prote
853	6	1.2	360	2	D69025	pleiotropic regula
854	6	1.2	360	2	S53536	MAP kinase (EC 2.7
855	6	1.2	360	2	T49066	cell division cycl
856	6	1.2	360	2	S12662	S-adenosylmethioni
857	6	1.2	360	2	S28047	TUB13 protein - po
858	6	1.2	360	2	A72423	D-mannosate hydrol
859	6	1.2	361	2	T01934	adenosylmethionine
860	6	1.2	361	2	S74433	GDP-D-mannose dehy
861	6	1.2	362	2	D81606	conserved hypotet
862	6	1.2	362	2	A12061	cit74 hypothetical
863	6	1.2	363	2	S23137	glycerol-3-phospha
864	6	1.2	363	2	T26385	hypothetical prote
865	6	1.2	363	2	S16998	polygalacturonase
866	6	1.2	363	2	S49222	adenosylmethionine
867	6	1.2	363	2	S43159	outer membrane por
868	6	1.2	363	2	UC1275	sensor protein bas
869	6	1.2	363	2	UX0285	hypothetical prote
870	6	1.2	363	2	T27335	probable membrane
871	6	1.2	364	2	T02801	Fe-binding protein
872	6	1.2	365	2	B54138	probable cell divi
873	6	1.2	365	2	A71005	hypothetical prote
874	6	1.2	365	2	F70988	tyrosine--tRNA lig
875	6	1.2	366	2	S75410	hypothetical prote
876	6	1.2	368	2	F72281	hypothetical prote
877	6	1.2	368	2	E71204	hypothetical prote
878	6	1.2	369	1	TYFVAF	transforming prote
879	6	1.2	369	1	B64921	conserved hypotet
880	6	1.2	369	1	H64611	probable transamin
881	6	1.2	369	2	G71902	variable major pro
882	6	1.2	369	2	S11960	conserved hypotet
883	6	1.2	369	2	T08266	hypothetical prote
884	6	1.2	370	2	C75043	hypothetical prote
885	6	1.2	370	2	T57555	c-Ha1 protein - mo
886	6	1.2	370	2	T45773	mucin-like protein
887	6	1.2	371	2	C64499	isocitrate dehydro
888	6	1.2	371	2	T45776	mucin-like protei
889	6	1.2	372	2	H81291	hypothetical protei
890	6	1.2	372	2	S24378	hypothetical prote
891	6	1.2	372	2	C64791	ybk protein - Esc
892	6	1.2	372	2	C70590	probable nict prot
893	6	1.2	373	2	A41918	transforming growt
894	6	1.2	373	2	G64482	hypothetical GTP-b
895	6	1.2	373	2	T10577	hypothetical prote
896	6	1.2	374	2	S76392	phosphoglycerate k
897	6	1.2	374	2	A41096	GTP-binding regula
898	6	1.2	374	2	T05891	myb-related protei
899	6	1.2	374	2	A57583	histone acetyltran
900	6	1.2	374	2	S46810	hypothetical prote
901	6	1.2	374	2	T34435	hypothetical prote
902	6	1.2	374	2	F81693	n1S protein, prob
903	6	1.2	375	2	A64398	hypothetical prote
904	6	1.2	376	2	C81272	probable amlnotran
905	6	1.2	376	2	T43846	translation elonga
906	6	1.2	377	2	I57490	guanine nucleotide
907	6	1.2	377	2	F69935	conserved hypotet
908	6	1.2	377	2	A35795	carbinol dehydrat
909	6	1.2	378	1	S17412	ubiquitinol--cytoch
910	6	1.2	378	2	S54056	probable membrane
911	6	1.2	379	2	E70937	hypothetical prote
912	6	1.2	379	2	I40873	serine proteinase
913	6	1.2	379	2	S42543	hypothetical prote
914	6	1.2	379	2	E72284	oxidoreductase, al
915	6	1.2	380	2	H71191	hypothetical prote
916	6	1.2	380	2	E69319	flavoprotein reduc
917	6	1.2	380	2	D44490	retrovirus-related
918	6	1.2	381	1	CMBO	chymosin (EC 3.4.2
919	6	1.2	381	1	CM5HB	chymosin (EC 3.4.2
920	6	1.2	381	2	E75458	acyl-CoA dehydroge
921	6	1.2	382	2	T14596	gag protein - alim
922	6	1.2	382	2	T46707	probable Iprn prot
923	6	1.2	384	2	A70805	conserved hypotet
924	6	1.2	384	2	T08318	syndecan-3 - chick
925	6	1.2	384	2	A44146	aldehyde dehydroge
926	6	1.2	385	2	JC5019	hypothetical prote
927	6	1.2	385	2	T19201	reverse transcript
928	6	1.2	385	2	S54987	reverse transcript
929	6	1.2	385	2	S54988	reverse transcript
930	6	1.2	385	2	S54990	reverse transcript
931	6	1.2	385	2	S54991	reverse transcript
932	6	1.2	385	2	S54992	reverse transcript
933	6	1.2	385	2	S54995	reverse transcript
934	6	1.2	385	2	S54997	reverse transcript
935	6	1.2	387	2	E75184	hypothetical prote
936	6	1.2	389	2	T22611	hypothetical prote
937	6	1.2	390	1	A42724	transcription init
938	6	1.2	390	1	TVMCB	transforming prote
939	6	1.2	390	2	H70353	ornithine decarbox
940	6	1.2	390	2	T39975	sarcosine oxidase
941	6	1.2	391	2	D64366	hypothetical prote
942	6	1.2	391	2	S17818	bacteriophage phi
943	6	1.2	391	2	B70608	probable PPE prote
944	6	1.2	392	2	T00669	polygalacturonase
945	6	1.2	392	2	T37173	probable cystathio
946	6	1.2	392	2	S11998	finger protein odd
947	6	1.2	393	2	T23602	hypothetical prote
948	6	1.2	393	2	E81691	conserved hypotet
949	6	1.2	394	2	S57099	hypothetical prote
950	6	1.2	395	2	A47701	hypothetical prote
951	6	1.2	395	2	B64338	aspartic proteinas
952	6	1.2	395	2	S44454	hypothetical prote
953	6	1.2	396	2	PS0075	transcription fact
954	6	1.2	396	2	S34963	myosin heavy chain
955	6	1.2	396	2	S26391	ribx protein - Shi
956	6	1.2	397	2	S30023	tail sheath protei
957	6	1.2	397	2	T00098	antiviral protein
958	6	1.2	398	2	A55581	hypothetical prote
959	6	1.2	398	2	T30013	nasc protein - Kle
960	6	1.2	398	2	S77529	hypothetical prote
961	6	1.2	398	2	T01921	hypothetical prote
962	6	1.2	399	2	T38388	translation elonga
963	6	1.2	400	1	A28172	hypothetical wd-40
964	6	1.2	400	1	A39822	spasmolysin precu
965	6	1.2	400	2	C69757	leukosialin precu
966	6	1.2	400	2	S70585	transporter homolo
967	6	1.2	400	2	S24759	nitrite reductase
968	6	1.2	400	2	A42727	probable transposa
969	6	1.2	400	2	G71900	hypothetical prote
970	6	1.2	401	2	E75213	n1S protein PAB01
971	6	1.2	401	2	T01001	hypothetical prote
972	6	1.2	401	2	H72765	probable flag endo
973	6	1.2	402	2	A70559	hypothetical prote
974	6	1.2	402	2	A70882	probable PPE prote
975	6	1.2	403	2	T45712	NAD-dependent mala
976	6	1.2	403	2	B81096	uracil permease NM
977	6	1.2	404	2	T25453	hypothetical prote
978	6	1.2	404	2	T34085	hypothetical prote

A:Cross-references: EMBL:X82629
 C:Genetics:
 A:Gene: GDB:MOX2
 A:Cross-references: GDB:120703; OMIM:155970
 A:Map position: 3q13-3q13
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:187-243/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0085;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHHHHHHHHH 12
 |||||
 Db 67 GHHHHHHHHH 77

RESULT 8
 B49122
 homeobox protein Mox-2 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 19-Dec-1993 #sequence_revision 17-Nov-1995 #text_change 24-Sep-1999
 C:Accession: B49122; S41779
 R:Candia, A.F.; Hu, J.; Crosby, J.; Lalle, P.A.; Nodén, D.; Nadeau, J.H.; Wright, C.V.
 Development 116, 1123-1136, 1992
 A>Title: Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed
 A:Reference number: A49122; MUID:93201999
 A:Accession: B49122
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-303 <CAN1>
 A:Cross-references: EMBL:Z16406; NID:957949; PIDN:CAA78899.1; PID:957950
 A:Experimental source: C57BL
 A>Note: the complete translation is not shown
 R:Candia, A.F.; Kovalik, J.P.; Wright, C.V.E.
 Nucleic Acids Res. 21, 4982, 1993
 A>Title: Amino acid sequence of Mox-2 and comparison to its Xenopus and rat homologs.
 A:Reference number: S41469; MUID:94232829
 A:Accession: S41779
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-303 <CAN2>
 A:Cross-references: EMBL:Z16406; NID:957949; PIDN:CAA78899.1; PID:957950
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:187-243/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0085;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHHHHHHHHH 12
 |||||
 Db 67 GHHHHHHHHH 77

RESULT 9
 A48130
 growth arrest-specific homeobox protein Cvx - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
 C:Accession: A48130; S31976
 R:Gorski, D.H.; Lepage, D.F.; Patel, C.V.; Copeland, N.G.; Jenkins, N.A.; Walsh, K.
 Mol. Cell. Biol. 13, 3722-3733, 1993
 A>Title: Molecular cloning of a diverged homeobox gene that is rapidly down-regulated during growth arrest
 A:Reference number: A48130; MUID:93268321
 A:Accession: A48130
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-303 <GOR>

A:Cross-references: GB:Z17223; NID:957951; PIDN:CAA78931.1; PID:957952
 A:Experimental source: aorta
 A>Note: sequence extracted from NCBI backbone (NCBIN:132842, NCBI:132843)
 R:Gorski, D.H.; Lepage, D.F.; Patel, C.V.; Copeland, N.G.; Jenkins, N.A.; Walsh, K.
 Submitted to the EMBL Data Library, October 1992
 A>Description: Cvx: A diverged homeodomain gene which is rapidly down-regulated follo
 A:Reference number: S31976
 A:Accession: S31976
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-182, 'S', 184-303 <GO2>
 A:Cross-references: EMBL:Z17223
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:187-243/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0085;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHHHHHHHHH 12
 |||||
 Db 67 GHHHHHHHHH 77

RESULT 10
 S39406
 homeotic protein otx1 - human
 C:Species: Homo sapiens (man)
 C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 15-Oct-1999
 C:Accession: S39406
 R:Strome, A.; Acampora, D.; Mallamaci, A.; Stornaiuolo, A.; d'Apice, M.R.; Nigro, V.
 EMBO J. 12, 2735-2747, 1993
 A>Title: A vertebrate gene related to orthodenticle contains a homeodomain of the bic
 A:Reference number: S35345; MUID:93327763
 A:Accession: S39406
 A>Status: Preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-354 <SIM>
 C:Genetics:
 A:Gene: otx1
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:39-95/Domain: homeobox homology <HOX>

Query Match 2.1%; Score 11; DB 2; Length 354;
 Best Local Similarity 100.0%; Pred. No. 0.0098;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHHHHHHHHH 12
 |||||
 Db 291 GHHHHHHHHH 301

RESULT 11
 S35345
 otx1 protein - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
 C:Accession: S35345
 R:Strome, A.; Acampora, D.; Mallamaci, A.; Stornaiuolo, A.; d'Apice, M.R.; Nigro, V.
 EMBO J. 12, 2735-2747, 1993
 A>Title: A vertebrate gene related to orthodenticle contains a homeodomain of the bic
 A:Reference number: S35345; MUID:93327763
 A:Accession: S35345
 A:Molecule type: DNA
 A:Residues: 1-355 <SIM>
 C:Genetics:
 A:Gene: otx1
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:39-95/Domain: homeobox homology <Hox>

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 355;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHNNHHNNH 12
|||||
DB 291 GHHNNHHNNH 301

RESULT 12

homeodomain protein ctx1 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999
C:Accession: I56547
R:Frantz, G.D.; Weimann, J.M.; Levin, M.E.; McConnell, S.K.
J. Neurosci. 14, 5725-5740, 1994
A:Title: Ctx1 and Ctx2 define layers and regions in developing cerebral cortex and cereb
A:Reference number: I56547; MUID:95016961
A:Accession: I56547
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: GB:I32602; NID:9535739; PIDN:AAA53557.1; PID:9535740
C:Genetics:
A:Gene: CTX1
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:39-95/Domain: homeobox homology <Hox>

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 355;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHNNHHNNH 12
|||||
DB 291 GHHNNHHNNH 301

RESULT 13

transcription repressor protein YY1 - human
N:Alternate names: transcription repressor/activator protein NF-E1
C:Species: Homo sapiens (man)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 08-Oct-1999
C:Accession: A40350; S78494; S33712; A56419
R:Shi, Y.; Seto, E.; Chang, L.S.; Shenk, T.
Cell 67, 377-388, 1991
A:Title: Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and re
A:Reference number: A40350; MUID:92005716
A:Accession: A40350
A:Molecule type: mRNA
A:Residues: 1-414 <SHI>
A:Cross-references: GB:M77698; NID:9186767; PIDN:AAA59467.1; PID:9186768
A:Experimental source: HeLa cells
A>Note: The authors translated the codon CGC for residue 371 as Lys, CGA for residue 375
R:Whitson, R.H.; Huang, T.; Dang, J.; Itakura, K.
Submitted to the EMBL Data Library, July 1992
A:Description: Observed and predicted DNA binding of a zinc finger protein which recogni
A:Reference number: S78494
A:Accession: S78494
A:Molecule type: mRNA
A:Residues: 1-195, G', 197-414 <WHI>
A:Cross-references: EMBL:Z14077; NID:938010; PIDN:CAA78455.1; PID:938011
R:Park, K.; Atchison, M.L.
Proc. Natl. Acad. Sci. U.S.A. 88, 9804-9808, 1991
A:Title: Isolation of a candidate repressor/activator, NF-E1 (YY-1), delta, that binds t
A:Reference number: A56419; MUID:92052179
A:Accession: S33712

A:Molecule type: mRNA
A:Residues: 1-64, R', 66-195, G', 197-414 <PAR>
A:Cross-references: GB:M76541; NID:9189173; PIDN:AAA59926.1; PID:9189174
C:Keywords: DNA binding; transcription regulation; zinc finger
F:298-330/Region: zinc finger CCH motif
F:327-347/Region: zinc finger CCH motif
F:355-377/Region: zinc finger CCH motif
F:385-407/Region: zinc finger CCH motif

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 414;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHNNHHNNH 12
|||||
DB 69 GHHNNHHNNH 79

RESULT 14

delta/YY1/NF-E1/UCBP transcription factor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: A48273; A42055
R:Saitany, G.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 90, 5559-5563, 1993
A:Title: Characterization of the mouse gene that encodes the delta/YY1/NF-E1/UCBP tr
A:Reference number: A48273; MUID:93296177
A:Accession: A48273
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-414 <RES>
A:Cross-references: GB:LI3968; NID:9293847; PIDN:AAA40477.1; PID:9293849
R:Flanagan, J.R.; Becker, K.G.; Ennlist, D.L.; Gleason, S.L.; Driggers, P.R.; Levi, B.
Mol. Cell. Biol. 12, 38-44, 1992
A:Title: Cloning of a negative transcription factor that binds to the upstream conser
A:Reference number: A42055; MUID:92107191
A:Accession: A42055
A:Molecule type: mRNA
A:Residues: 1-414 <FLA>
A:Cross-references: GB:M73963; NID:9202270; PIDN:AAA40522.1; PID:9202271
A>Note: sequence extracted from NCBI backbone (NCBIN:74641, NCBI:P:74642)
C:Genetics:
A:Introns: 227/1; 281/2; 301/3; 354/3

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 414;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHNNHHNNH 12
|||||
DB 70 GHHNNHHNNH 80

RESULT 15

transcription factor delta - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 05-Nov-1999
C:Accession: A56418
R:Hartharan, N.; Kelley, D.E.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 88, 9799-9803, 1991
A:Title: delta, a transcription factor that binds to downstream elements in several p
A:Reference number: A56418; MUID:92052178
A:Accession: A56418
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-414 <HAR>
A:Cross-references: GB:M74590; NID:9192940; PIDN:AAA37521.1; PID:9192941
C:Keywords: transcription factor; zinc finger

Query Match 2.18; Score 11; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GHHHHHHHHH 12
Db 70 GHHHHHHHHH 80

Search completed: August 22, 2000, 17:43:27
Job time: 1131 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:42:01 ; Search time 43.07 Seconds
(without alignments)
374.863 Million cell updates/sec

Title: US-08-699-716a-2

Perfect score: 521
Sequence: 1 MGNHHHHHHSSGHIDDD.....RFIQKDYDVMQRLDDITSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 85661 seqs, 30989116 residues

Word size : 0

Total number of hits satisfying chosen parameters: 85661

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	36.7	326	1	LCRV_YERPE
2	170	32.6	170	1	CARL_YERPE
3	82	15.7	326	1	LCRV_YERPS
4	11	2.1	303	1	MOX2_HUMAN
5	11	2.1	303	1	MOX2_MOUSE
6	11	2.1	303	1	MOX2_MOUSE
7	11	2.1	354	1	OTX1_HUMAN
8	11	2.1	355	1	OTX1_MOUSE
9	11	2.1	355	1	OTX1_MOUSE
10	11	2.1	414	1	TYR1_HUMAN
11	11	2.1	414	1	TYR1_MOUSE
12	11	2.1	466	1	ZIC3_MOUSE
13	11	2.1	467	1	ZIC3_HUMAN
14	11	2.1	611	1	SNF1_CANGA
15	11	2.1	619	1	SNF1_CANTR
16	10	1.9	298	1	MOX2_XENLA
17	10	1.9	331	1	HXAI_MOUSE
18	10	1.9	335	1	HXAI_HUMAN
19	10	1.9	351	1	HRPX_PLATO
20	10	1.9	375	1	CAR2_DICDI
21	10	1.9	387	1	QIN_AVIS3
22	10	1.9	390	1	PC_DROME
23	10	1.9	410	1	BR3B_HUMAN
24	10	1.9	411	1	BR3B_MOUSE
25	10	1.9	430	1	NOR2_RAT
26	10	1.9	449	1	GATE_HUMAN
27	10	1.9	451	1	BFI_CHICK
28	10	1.9	469	1	BFI_HUMAN
29	10	1.9	477	1	BFI_HUMAN
30	10	1.9	480	1	BFI_MOUSE
31	10	1.9	481	1	BFI_MOUSE
32	10	1.9	620	1	SNF1_CANAL
33	10	1.9	623	1	PNT1_DROME

34	10	1.9	626	1	NR43_HUMAN	Q92570 homo sapien
35	10	1.9	628	1	NR43_RAT	P51179 rattus norv
36	10	1.9	633	1	SNF1_YEAST	P06782 saccharomyc
37	10	1.9	657	1	KNOB_PLAEN	P06719 plasmodium
38	10	1.9	665	1	LAMA_XENLA	P11048 xenopus lae
39	10	1.9	717	1	UBPB_YEAST	P36026 saccharomyc
40	10	1.9	763	1	DYRK_HUMAN	Q13627 homo sapien
41	10	1.9	763	1	DYRK_MOUSE	O61214 mus musculu
42	10	1.9	763	1	DYRK_RAT	O63470 rattus norv
43	10	1.9	950	1	URB1_DROME	P40349 ussiliago ma
44	10	1.9	1312	1	PIPL_DROME	P23455 drosophila
45	10	1.9	1356	1	ROM2_YEAST	P51862 saccharomyc
46	10	1.9	2164	1	CCAA_MOUSE	P97445 mus musculu
47	10	1.9	2212	1	CCAA_RAT	P54282 rattus norv
48	10	1.9	2505	1	CCAA_HUMAN	O00555 homo sapien
49	10	1.9	143	1	Y970_METJA	O58380 methanococc
50	9	1.7	332	1	DLX2_MOUSE	P40764 mus musculu
51	9	1.7	333	1	HXAI_RAT	O08656 rattus norv
52	9	1.7	359	1	HNKI_MAIZE	P24345 zea mays (m
53	9	1.7	421	1	BR3A_MOUSE	P17208 mus musculu
54	9	1.7	423	1	BR3A_HUMAN	O01851 homo sapien
55	9	1.7	428	1	FKR4_MOUSE	O64733 mus musculu
56	9	1.7	469	1	A2AC_DIDMA	P35405 didelphis m
57	9	1.7	473	1	KNOB_PLAFA	P33405 plasmodium
58	9	1.7	568	1	DISC_DROME	P13817 plasmodium
59	9	1.7	597	1	STAR_DROME	P23792 drosophila
60	9	1.7	634	1	KNOB_PLARG	P42519 drosophila
61	9	1.7	2044	1	CCAH_HUMAN	P09346 plasmodium
62	9	1.7	2254	1	CCAG_RAT	O95180 homo sapien
63	9	1.7	2424	1	CCAA_RABIT	O54898 rattus norv
64	9	1.5	59	1	HPN_HELPY	P27884 oryctolagus
65	8	1.5	323	1	MAF1_MOUSE	O48251 helicobacte
66	8	1.5	323	1	MAF1_RAT	P54841 mus musculu
67	8	1.5	323	1	OTX1_BARE	P54842 rattus norv
68	8	1.5	351	1	CAY2_CAEEL	O91994 brachydanio
69	8	1.5	409	1	BR11_BARE	O18879 caenorhabdi
70	8	1.5	437	1	BP12_BARE	O90436 brachydanio
71	8	1.5	486	1	MEC2_HUMAN	P56224 brachydanio
72	8	1.5	495	1	BRN1_MOUSE	P51608 homo sapien
73	8	1.5	497	1	BRN1_RAT	P31361 mus musculu
74	8	1.5	500	1	BRN1_HUMAN	O63262 rattus norv
75	8	1.5	508	1	CROC_DROME	P20264 homo sapien
76	8	1.5	538	1	ZIC2_MOUSE	P33027 drosophila
77	8	1.5	609	1	OPA_DROME	O63520 mus musculu
78	8	1.5	684	1	EPB4_HCVVA	P39768 drosophila
79	8	1.5	972	1	TOP1_DROME	P17151 human cytom
80	8	1.5	1133	1	ATX9_TETTH	P30189 drosophila
81	8	1.5	1597	1	SOL_DROME	O95050 teletaymena
82	8	1.3	16	1	LPH1_ECOLI	P27398 drosophila
83	7	1.3	91	1	S10B_BOVIN	P03058 escherichia
84	7	1.3	91	1	S10B_HUMAN	P02638 bos taurus
85	7	1.3	91	1	S10B_MOUSE	P04271 homo sapien
86	7	1.3	91	1	S10B_RAT	P04631 mus musculu
87	7	1.3	97	1	VE7_HPV23	P04631 rattus norv
88	7	1.3	100	1	VE7_HPV38	P050781 human papil
89	7	1.3	104	1	PTLA_STRW	O80908 human papil
90	7	1.3	105	1	PTLA_LACIA	P26426 streptococc
91	7	1.3	111	1	CB21_CARPI	P23532 lactococcus
92	7	1.3	114	1	ASR2_LYCES	P38582 carnobacter
93	7	1.3	127	1	RL21_SYNP7	P37219 lycopersico
94	7	1.3	143	1	MOCR_RHISN	O94385 synchrococc
95	7	1.3	143	1	YLAB_ENTHS	P55363 rhizobium s
96	7	1.3	147	1	FER5_METJA	O34913 enterococcu
97	7	1.3	153	1	ENRP_BPP21	O57669 methanococc
98	7	1.3	162	1	RN_BACIN	P27358 bacteriophag
99	7	1.3	162	1	RN_BACPU	P27358 bacteriophag
100	7	1.3	182	1	DYR_DROME	P00649 bacillus in
101	7	1.3	193	1	YDCM_HAEIN	P48068 bacillus pu
102	7	1.3	216	1	UREY_RHIME	P71719 drosophila
103	7	1.3	252	1	COR2_EMENT	P73396 haemophilus
104	7	1.3	253	1	KDSB_HAEIN	P43880 haemophilus
105	7	1.3	297	1	CTAA_BACET	P13588 emricella
106	7	1.3	302	1	HYPB_BRAJA	P44490 haemophilus
						O04443 bacillus fi
						O45257 bradyrhizob

107	7	1.3	321	1	CANC_T0BAC	P27141	180	1.2	65	1	TRBK_RHISN	P55401
108	7	1.3	328	1	DLX2_HUMAN	007687	181	1.2	70	1	ATP1_BOVIN	Q00361
109	7	1.3	328	1	OPY1_YEAST	P38271	182	1.2	78	1	KCRU_PIG	Q29577
110	7	1.3	329	1	CAHX_FLAPR	P46281	183	1.2	80	1	COXK_MOUSE	P56992
111	7	1.3	330	1	CAH1_FLALI	P46512	184	1.2	80	1	PSY1_SYNP2	P31662
112	7	1.3	330	1	CAHX_FLABR	P46511	185	1.2	82	1	BY1M_HERRA	P25777
113	7	1.3	331	1	CAHX_FLABI	P46510	186	1.2	82	1	YE1M_HERRA	P25781
114	7	1.3	332	1	G3P3_ECOLI	P33898	187	1.2	85	1	YOR4_BPSPP	Q38440
115	7	1.3	336	1	GUNA_RUMFL	P16189	188	1.2	89	1	TUBE_MYCTU	P02944
116	7	1.3	348	1	YHHT_HAEIN	P46466	189	1.2	92	1	RPC1_BPP22	P03461
117	7	1.3	350	1	HMRO_DROME	P10181	190	1.2	92	1	Y766_RICPR	O926b5
118	7	1.3	350	1	Y098_SYNY3	O55880	191	1.2	93	1	VE7_HPV50	O80828
119	7	1.3	359	1	Y0R3_CAREL	O09586	192	1.2	95	1	NO1E_RHILP	P23717
120	7	1.3	362	1	VM25_BORHE	P32779	193	1.2	99	1	B2WG_RABIT	P01885
121	7	1.3	364	1	VM21_BORHE	P21875	194	1.2	100	1	VE7_HPV22	P50780
122	7	1.3	374	1	YDVB_SCHPO	O14229	195	1.2	101	1	VG7_BPP03	Q37887
123	7	1.3	374	1	YMP4_CAREL	O10948	196	1.2	103	1	PTLA_STRAU	P02909
124	7	1.3	376	1	T2A4_HUMAN	P52655	197	1.2	106	1	Y115_ADE07	P03288
125	7	1.3	380	1	CAH2_CHLRE	P24258	198	1.2	107	1	SSI1_STRGI	P28392
126	7	1.3	385	1	SELP_RAT	P25236	199	1.2	111	1	YCX1_CHLRE	P05723
127	7	1.3	392	1	PURT_PASHA	P46927	200	1.2	112	1	VPX_HVZD1	P17760
128	7	1.3	398	1	TBX1_HUMAN	O43435	201	1.2	112	1	VPX_HVZG1	P18045
129	7	1.3	401	1	PRRB_ECOLI	P17222	202	1.2	112	1	Y13K_SSV1	P20920
130	7	1.3	408	1	FRF2_HUMAN	Q12947	203	1.2	113	1	PSAK_CHLRE	P14225
131	7	1.3	410	1	T1S1_ECOLI	P10465	204	1.2	113	1	VC21_VACCC	P21105
132	7	1.3	425	1	POU1_BRARE	P31366	205	1.2	113	1	VP21_HV2BE	P18099
133	7	1.3	441	1	GAT6_RAT	P46153	206	1.2	115	1	KG20_PYRAB	P56813
134	7	1.3	443	1	ZP23_BRARE	P79745	207	1.2	121	1	KDGL_ECOLI	P00556
135	7	1.3	444	1	GAT6_MOUSE	O61169	208	1.2	121	1	LCA_MACRG	P07458
136	7	1.3	485	1	CG2A_CAREL	P34638	209	1.2	122	1	G5TB_BACSU	P26307
137	7	1.3	480	1	HMS1_YEAST	P54725	210	1.2	125	1	YOR3_YEREN	P28836
138	7	1.3	492	1	MEC2_RAT	O00566	211	1.2	127	1	YQJB_ECOLI	P97885
139	7	1.3	515	1	PEZ2_SCHPO	P78968	212	1.2	130	1	S205_RAT	O48405
140	7	1.3	518	1	YLM6_CAREL	P34409	213	1.2	131	1	GP51_BPSPL	O67469
141	7	1.3	519	1	SAZD_HUMAN	Q12788	214	1.2	131	1	YE98_AQUAE	O25754
142	7	1.3	532	1	VLI_HPV03	P36731	215	1.2	133	1	EXD1_HELPY	P07088
143	7	1.3	535	1	DPPA_ECOLI	P23847	216	1.2	134	1	ACP_BRACM	P03786
144	7	1.3	570	1	NU5M_PARTE	P15584	217	1.2	135	1	V47_BPP7	O06555
145	7	1.3	572	1	PT1_STAUV	P51183	218	1.2	140	1	LCA_MACEU	P49324
146	7	1.3	579	1	G160_HUMAN	O08378	219	1.2	140	1	PRO3_WHEAT	P16138
147	7	1.3	587	1	HE_PARI1	P22757	220	1.2	140	1	VHM3_CAVPK	P49333
148	7	1.3	604	1	Y1J3_YEAST	P47030	221	1.2	141	1	PRO2_WHEAT	P24353
149	7	1.3	615	1	ALP1_YEAST	P46880	222	1.2	142	1	RR8_ASTLO	O56069
150	7	1.3	617	1	YA42_HAEIN	O57175	223	1.2	144	1	MARR_SALTY	P22446
151	7	1.3	627	1	CIAT_CAREL	P32756	224	1.2	144	1	VGR_BPP1	O07906
152	7	1.3	682	1	SNK_MOUSE	P53351	225	1.2	144	1	YRA1_BACSU	O19825
153	7	1.3	746	1	TDCE_ECOLI	P42632	226	1.2	148	1	YRA1_CAREL	O04837
154	7	1.3	785	1	SECA_MYCLE	O32932	227	1.2	148	1	SSB_HUMAN	P28042
155	7	1.3	785	1	YG51_YEAST	P50089	228	1.2	151	1	SSB_RAT	P36933
156	7	1.3	808	1	SEA2_MYCTU	O50612	229	1.2	155	1	VPR_BPP2	P52146
157	7	1.3	817	1	HUNB_MUSDO	O01778	230	1.2	158	1	GREK_ECOLI	O54143
158	7	1.3	836	1	PDC2_CANAL	O60035	231	1.2	159	1	SMRA_TREHY	P17900
159	7	1.3	886	1	MCM2_CANAL	P55861	232	1.2	161	1	UREA_PROMI	P11390
160	7	1.3	887	1	MCM2_XENLA	P49735	233	1.2	162	1	PHAC_STNPG	Q10795
161	7	1.3	895	1	MCM2_HUMAN	P49736	234	1.2	162	1	RS16_MYCTU	O78478
162	7	1.3	904	1	MCM2_DROME	P27610	235	1.2	163	1	ATPX_GUTH	P33661
163	7	1.3	906	1	CB1_HUMAN	P22681	236	1.2	163	1	GVH2_HALSA	P33661
164	7	1.3	906	1	RPOL_BPKT1	P18147	237	1.2	164	1	LSPK_ECOLI	P00804
165	7	1.3	944	1	MMU3_MYCTU	O53657	238	1.2	165	1	LSPK_ECOLI	P13314
166	7	1.3	946	1	YB26_YEAST	P38280	239	1.2	166	1	RNH2_LACLC	O30415
167	7	1.3	946	1	YB26_YEAST	P40559	240	1.2	166	1	YNI1_CAREL	P45561
168	7	1.3	946	1	YB26_YEAST	P40559	241	1.2	169	1	BENB_ACICL	P07770
169	7	1.3	1356	1	HEH1_PODAN	O00808	242	1.2	169	1	GBP_XENLA	O93343
170	7	1.3	1705	1	CYAA_BORBE	O57506	243	1.2	170	1	GPV2_HALSA	P17942
171	7	1.3	1706	1	CYAA_BORBE	O57506	244	1.2	170	1	LSPK_PEEFL	O32418
172	7	1.3	1912	1	PTPD_HUMAN	P23468	245	1.2	172	1	APT_STAUV	P08407
173	7	1.3	2326	1	CCAB_DISOM	P56698	246	1.2	173	1	PRSE_ECOLI	P42186
174	6	1.2	2483	1	PCX_DROME	P18450	247	1.2	173	1	PRSE_ECOLI	P47939
175	6	1.2	16	1	TRYP_FELCA	P81071	248	1.2	175	1	RP18_MOUSE	P37981
176	6	1.2	31	1	PSBT_CHLRE	P81071	249	1.2	179	1	IPYR_THEAC	P12977
177	6	1.2	35	1	PSBT_CHLVU	P56327	250	1.2	179	1	RLS_BACSU	P71378
178	6	1.2	54	1	IOVO_CASCA	P05559	251	1.2	182	1	CYPA_MYCTU	P04261
179	6	1.2	54	1	IOVO_DRONO	P05560	252	1.2	182	1	K2C3_BOVIN	

253	6	1.2	182	1	K2C4_BOVIN	P04260	bos taurus
254	6	1.2	184	1	P12_LUPPO	P16148	lupinus:pol
255	6	1.2	184	1	TX1_MOUSE	P70323	mus musculus
256	6	1.2	185	1	NIR_LEPMC	P45004	leptosphaer
257	6	1.2	192	1	PABA_STRUT	P27637	streptomyc
258	6	1.2	192	1	YS10_BORBU	P70837	botryella bu
259	6	1.2	193	1	Y983_HAETN	P43907	haemophilus
260	6	1.2	194	1	RS7_ARCFU	O23866	archaeoglob
261	6	1.2	195	1	INT_CAPHI	P28171	capra hircu
262	6	1.2	197	1	YGVV_ECOLI	P50061	escherichia
263	6	1.2	198	1	APPF_MYCGA	P33256	mycoplasma
264	6	1.2	198	1	FRS6_RHORA	P43378	rhodococcus
265	6	1.2	199	1	YCXI_PORPU	P51354	porphyra pu
266	6	1.2	199	1	PA10_YEAST	P48363	saccharomyc
267	6	1.2	201	1	YAMC_SCHPO	O10186	schizosacch
268	6	1.2	204	1	YAS7_ACTAC	O51728	actinobacil
269	6	1.2	205	1	RNH2_ARCRU	O26634	archaeoglob
270	6	1.2	208	1	Y4IN_RHISN	P53497	rhizobium s
271	6	1.2	209	1	GC_NETGO	O87406	neisseria g
272	6	1.2	210	1	CD8B_PONPY	P30434	pongo pygma
273	6	1.2	210	1	YG21_YEAST	P53251	saccharomyc
274	6	1.2	212	1	PECF_MASIA	P28730	maestigoclad
275	6	1.2	212	1	KRRC_ARATH	O94W17	arabidopsis
276	6	1.2	213	1	HIS1_BACSU	O34520	bacillus su
277	6	1.2	213	1	HPRT_MUSSP	O64531	mus spretus
278	6	1.2	217	1	HPRT_CRIGR	P00494	cricetulus
279	6	1.2	217	1	HPRT_MERUN	P47959	meriones un
280	6	1.2	217	1	HPRT_MOUSE	P00493	mus musculus
281	6	1.2	218	1	HPRT_RAT	P27605	rattus norv
282	6	1.2	219	1	THIE_HELPY	O23514	helicobacte
283	6	1.2	220	1	DEOC_MYCPI	P47722	mycoplasma
284	6	1.2	220	1	YGCA_VIBSS	P55136	vibrio sp.
285	6	1.2	221	1	CART_MOUSE	O08331	mus musculus
286	6	1.2	223	1	YEL2_EBV	P03203	epstein-bar
287	6	1.2	224	1	KTHX_CHICK	P04047	gallus galli
288	6	1.2	226	1	Y703_METUA	O58114	methanococc
289	6	1.2	230	1	FLAH_METYO	O06641	methanococc
290	6	1.2	231	1	PSAF_SPIOL	P12355	spinacia ol
291	6	1.2	233	1	H11_GLYXA	P40263	glyptotendi
292	6	1.2	233	1	H11_GLYXA	P40266	glyptotendi
293	6	1.2	234	1	RK1_GUTTH	O78413	guillardia
294	6	1.2	234	1	Y61B_ECOLI	P24195	escherichia
295	6	1.2	239	1	TRPA_THEMA	P50908	thermotoga
296	6	1.2	240	1	YB20_PLAFA	O96178	plasmidium
297	6	1.2	243	1	IP72_AGR7	P06524	agrobacteri
298	6	1.2	243	1	IPR_AGR7	P14011	agrobacteri
299	6	1.2	243	1	YHFR_ECOLI	P45544	escherichia
300	6	1.2	244	1	TONB_VIBCH	O50442	vibrio chol
301	6	1.2	246	1	TRY1_CANFA	P06871	canis famli
302	6	1.2	247	1	ATP6_ACACA	O37385	acanthamoeb
303	6	1.2	248	1	Y1PA_YEAST	P53039	saccharomyc
304	6	1.2	249	1	YCIT_ECOLI	P70034	escherichia
305	6	1.2	250	1	Y035_METUA	O60350	methanococc
306	6	1.2	252	1	DEOR_ECOLI	P06217	escherichia
307	6	1.2	252	1	ONCM_HUMAN	P13725	homo sapien
308	6	1.2	252	1	TRPC_THEMA	O56319	thermotoga
309	6	1.2	252	1	YR2_CAEEL	O10006	caenorhabdi
310	6	1.2	253	1	UPPS_ECOLI	O47675	escherichia
311	6	1.2	253	1	SSRA_ARATH	P44434	arabidopsis
312	6	1.2	253	1	YTBQ_BACSU	P53560	bacillus su
313	6	1.2	253	1	YX04_CAEEL	O11111	caenorhabdi
314	6	1.2	255	1	MYB3_MAZE	P20025	zee mays (m
315	6	1.2	256	1	YKHI_YEAST	P36086	saccharomyc
316	6	1.2	257	1	ETXA_STAUD	P13163	staphylococ
317	6	1.2	258	1	STV_STRCO	O06851	streptomyc
318	6	1.2	258	1	Y098_MYCTU	O10859	mycobacteri
319	6	1.2	261	1	VP31_FRGV	P18178	frog virus
320	6	1.2	264	1	FLGG_BACSU	P23446	bacillus su
321	6	1.2	264	1	YUPZ_ECOLI	P33308	escherichia
322	6	1.2	267	1	Y617_SYNY3	O55707	synechocyst
323	6	1.2	268	1	CEBD_MOUSE	O00322	mus musculus
324	6	1.2	268	1	CEBD_RAT	O03484	rattus norv
325	6	1.2	268	1	FLHP_BACSU	P33753	bacillus su
326	6	1.2	269	1	CEBD_HUMAN	P49716	homo sapien
327	6	1.2	271	1	CART_RAT	P22676	homo sapien
328	6	1.2	271	1	HMR_XENLA	P47778	rattus norv
329	6	1.2	271	1	HMR_XENLA	P14887	xenopus lae
330	6	1.2	271	1	HXA5_MOUSE	P09631	mus musculus
331	6	1.2	271	1	Y1UB_VIBCH	O56646	vibrio chol
332	6	1.2	271	1	YW34_MYCTU	O05879	mycobacteri
333	6	1.2	272	1	DKH1_STRVN	P16542	streptomyc
334	6	1.2	272	1	HXA9_HUMAN	P31269	homo sapien
335	6	1.2	273	1	FLJM_CACR	O52529	caulobacter
336	6	1.2	273	1	IACG_AGRD	P29884	agrobacteri
337	6	1.2	273	1	T2C2_HERAD	P25259	herpetosiph
338	6	1.2	273	1	Y133_METUA	O57537	methanococc
339	6	1.2	274	1	FLJO_CACR	O52531	caulobacter
340	6	1.2	274	1	T2B1_HERAD	P25257	herpetosiph
341	6	1.2	274	1	YOHM_ECOLI	P25260	herpetosiph
342	6	1.2	274	1	YOHM_ECOLI	P76425	escherichia
343	6	1.2	276	1	BACH_HALHP	O48315	halobacteri
344	6	1.2	280	1	YME1_YEAST	O34433	saccharomyc
345	6	1.2	280	1	CEBE_HUMAN	O15744	homo sapien
346	6	1.2	281	1	NS2_MYCTU	O58484	methanococc
347	6	1.2	282	1	AROE_METUA	O10796	mycobacteri
348	6	1.2	287	1	DCOP_MYXXA	P44230	mycobacteri
349	6	1.2	288	1	PIP_BACCO	P46541	bacillus co
350	6	1.2	288	1	YIGM_ECOLI	P71849	escherichia
351	6	1.2	290	1	ARY3_MOUSE	P50266	mus musculus
352	6	1.2	291	1	GDBB_WHEAT	P06659	titlicum ae
353	6	1.2	291	1	NLA_DROME	O94248	dirosophila
354	6	1.2	292	1	YQFU_BACSU	P74478	bacillus su
355	6	1.2	293	1	RS2_MYCPN	P75560	mycoplasma
356	6	1.2	294	1	MED6_YEAST	P87872	saccharomyc
357	6	1.2	295	1	PERX_BRARA	P00434	brassica ra
358	6	1.2	296	1	YD28_METUA	O58724	methanococc
359	6	1.2	298	1	DAPA_RHIME	O59762	rhizobium m
360	6	1.2	300	1	APBA_PYRHO	O50098	pyrococcus
361	6	1.2	301	1	P061_HUMAN	O14868	homo sapien
362	6	1.2	301	1	P061_MOUSE	O07916	mus musculus
363	6	1.2	301	1	P061_MOUSE	P56223	rattus norv
364	6	1.2	301	1	P061_MOUSE	P12522	titlicum ae
365	6	1.2	302	1	GDBX_WHEAT	O50559	methanobact
366	6	1.2	303	1	KIME_MERTH	P12268	sus scrofa
367	6	1.2	303	1	OSTP_PIG	O14236	saccharomyc
368	6	1.2	303	1	PMG3_YEAST	P24280	saccharomyc
369	6	1.2	303	1	SC14_YEAST	O44104	dirosophila
370	6	1.2	304	1	G3P2_DROPS	O44105	dirosophila
371	6	1.2	304	1	G3P2_DROPS	P31316	mus musculus
372	6	1.2	305	1	GS82_MOUSE	P53389	rhizobium s
373	6	1.2	305	1	Y4CG_RHISN	O45008	haemophilus
374	6	1.2	306	1	OPPB_HAETN	P25668	pseudomonas
375	6	1.2	306	1	SDSB_PSES9	P75758	escherichia
376	6	1.2	307	1	YKIC_ECOLI	P75758	escherichia
377	6	1.2	307	1	YK55_YEAST	P61615	saccharomyc
378	6	1.2	310	1	VE2_HPVS6	P67678	human papil
379	6	1.2	311	1	PMG2_YEAST	O12608	saccharomyc
380	6	1.2	312	1	TRAD_HUMAN	O15528	homo sapien
381	6	1.2	314	1	ETFA_BRARA	P53373	bradyrhizob
382	6	1.2	314	1	YN03_YEAST	P33845	saccharomyc
383	6	1.2	316	1	PER1_ARAHY	P22155	atachalis hyp
384	6	1.2	320	1	THFR_SYNP7	P27477	synechococc
385	6	1.2	322	1	ARGC_ANASP	P44894	anaeana sp
386	6	1.2	323	1	ANX5_CYNPY	P70075	cynops pyr
387	6	1.2	323	1	ARGI_EMINI	O12611	emeritcella
388	6	1.2	323	1	CCG2_HUMAN	O59689	homo sapien
389	6	1.2	323	1	CCG2_MOUSE	O88602	mus musculus
390	6	1.2	323	1	ISPB_ECOLI	P19641	escherichia
391	6	1.2	325	1	Y0XA_BACSU	P39840	bacillus su
392	6	1.2	328	1	T2EB_YEAST	P61645	saccharomyc
393	6	1.2	328	1	Y36K_HALSP	P14321	halobacteri
394	6	1.2	332	1	BIOB_BACSH	P19206	bacillus sp
395	6	1.2	332	1	G3P1_DROME	P07466	dirosophila
396	6	1.2	332	1	G3P2_DROME	P07467	dirosophila
397	6	1.2	332	1	G3P_DROXY	O01597	dirosophila
398	6	1.2	332	1	YG2H_YEAST	P53250	saccharomyc

399	6	1.2	333	1	MP92_POAPR	P22885	poa piteus	472	6	1.2	377	1	GB13_HUMAN	Q14344	homo sapien
400	6	1.2	334	1	G3P1_ANAVA	P34916	anabaena va	473	6	1.2	377	1	YPER_BACSV	P54169	baicillus su
401	6	1.2	334	1	VAS1_VACCC	P21069	vaccinia v1	474	6	1.2	378	1	CYB_ELEMA	CYB8	elephas ma
402	6	1.2	334	1	VAS1_VACCV	Q01419	vaccinia v1	475	6	1.2	378	1	MNT3_CANAL	P87207	candida alb
403	6	1.2	334	1	VAS1_VARV	P33858	varicola vlr	476	6	1.2	378	1	YN9B_YEAST	P55923	saccharomyc
404	6	1.2	336	1	Y883_METJA	058933	methanococ	477	6	1.2	379	1	CYB_LOXAP	P24958	toxodonta a
405	6	1.2	338	1	DCUP_AOUAE	066667	aquifex aeo	478	6	1.2	379	1	HMB1_SOYBN	P46608	glycine max
406	6	1.2	338	1	SIFA_AOUAE	067087	aquifex aeo	479	6	1.2	379	1	YAXO_RHISN	P55707	rhizobium s
407	6	1.2	338	1	TRRS_ECOLI	P03837	escherichia	480	6	1.2	381	1	CHYM_BOVIN	P00794	bos taurus
408	6	1.2	339	1	SPSG_BACSV	P59627	baicillus su	481	6	1.2	381	1	CHYM_SHEEP	P18272	ovis aries
409	6	1.2	339	1	SRPA_SYNP7	055025	synchococ	482	6	1.2	383	1	PO13_NASVT	003271	onosaia vit
410	6	1.2	340	1	YLS1_YEAST	Q12524	saccharomyc	483	6	1.2	385	1	DH48_HUMAN	P48448	homo sapien
411	6	1.2	341	1	ETRA_SCHPO	P18790	saccharomyc	484	6	1.2	390	1	CB1_MLVCN	P23092	cas-n8-1 mu
412	6	1.2	342	1	RECA_ERMCA	P26344	erwinia car	485	6	1.2	390	1	RPSD_ANASP	P26683	anabaena sp
413	6	1.2	344	1	MURG_AOUAE	067238	aquifex aeo	486	6	1.2	390	1	SAOX_BACBP	P40859	baicillus sp
414	6	1.2	344	1	TMRA_TREPH	P29721	treponema p	487	6	1.2	391	1	BCHP_RHOCA	P26172	rhodobacter
415	6	1.2	345	1	VP10_MTV	P13093	wound tumor	488	6	1.2	391	1	Y532_METJA	057952	methanococ
416	6	1.2	347	1	F16P_YEAST	P09201	saccharomyc	489	6	1.2	392	1	CYSA_STRO	059829	methanococ
417	6	1.2	347	1	IDH_METJA	058991	methanococ	490	6	1.2	392	1	ODD_DROME	P23803	strepomyc
418	6	1.2	348	1	RMUB_STRMU	P95780	streplococ	491	6	1.2	394	1	YJ51_YEAST	P47127	saccharomyc
419	6	1.2	349	1	ASG2_HAEIN	P43843	haemophilus	492	6	1.2	395	1	YF01_BPP2	P22501	bacterioph
420	6	1.2	349	1	GPDA_DROAE	Q27556	deschiphila	493	6	1.2	395	1	Y305_METJA	057753	methanococ
421	6	1.2	349	1	YGB0_ECOLI	057261	escherichia	494	6	1.2	395	1	REBX_SHIDY	003583	shigella dy
422	6	1.2	352	1	GBA1_CRYPA	000580	cryphonectr	495	6	1.2	396	1	SP11_HUMAN	Q935K1	homo sapien
423	6	1.2	352	1	GPDA_DROVI	P07335	drosophila	496	6	1.2	397	1	SK18_YEAST	Q02793	saccharomyc
424	6	1.2	352	1	KAPC_DROME	P12370	drosophila	497	6	1.2	398	1	SME_SYNY3	P73345	synchocyst
425	6	1.2	352	1	Z185_MOUSE	Q62394	mus musculu	498	6	1.2	400	1	LEUK_HUMAN	P16150	homo sapien
426	6	1.2	354	1	PON2_MOUSE	Q62086	mus musculu	499	6	1.2	400	1	MUAI_XENLA	P10667	xenopus lae
427	6	1.2	354	1	RECA_PROMI	P11406	proteus mir	500	6	1.2	400	1	TRA3_RHIME	P80011	rhizobium m
428	6	1.2	354	1	YMA3_MYCBO	002279	mycobacteri	501	6	1.2	400	1	YAPO_RHISN	P55620	rhizobium s
429	6	1.2	355	1	AMIE_STRPN	P18765	streplococ	502	6	1.2	400	1	YCEI_BACSV	P56620	baicillus su
430	6	1.2	355	1	CMG2_SCHPO	P32434	schizosacch	503	6	1.2	401	1	P39_BRUAB	034691	baicillus su
431	6	1.2	355	1	LEB9_HUMAN	Q00182	homo sapien	504	6	1.2	401	1	PGK_SYNY3	006875	brucella ab
432	6	1.2	355	1	MTM2_METTF	P29568	methanobact	505	6	1.2	403	1	MHTP_ECOLI	P74421	synchocyst
433	6	1.2	358	1	G3PG_TRYBB	P22512	trypanosoma	506	6	1.2	404	1	MEIK_ECOLI	P77589	escherichia
434	6	1.2	358	1	RECA_XENBY	P96185	xenorhabdus	507	6	1.2	404	1	TY3H_CAEEL	P50303	caenorhabdi
435	6	1.2	358	1	SP11_MOUSE	09wtk8	mus musculu	508	6	1.2	405	1	PSB8_DROME	P90986	caenorhabdi
436	6	1.2	359	1	G3PG_TRYCR	P22513	trypanosoma	509	6	1.2	405	1	SDC3_CHICK	018413	drosophila
437	6	1.2	360	1	DCAM_SOLTU	Q04594	solanum tub	510	6	1.2	406	1	PSB8_HUMAN	P26261	galinus gall
438	6	1.2	360	1	MK14_HUMAN	Q16539	h. mtogen-a	511	6	1.2	406	1	PRB8_MOUSE	P47210	homo sapien
439	6	1.2	360	1	MK14_MOUSE	P47811	mus musculu	512	6	1.2	407	1	COAT_FHV	P52915	mus musculu
440	6	1.2	360	1	MK14_MOUSE	P70618	rattus norv	513	6	1.2	407	1	VG10_HSVSA	P18870	flock house
441	6	1.2	360	1	UXRA_THEMA	09wxs4	thermotoga	514	6	1.2	409	1	EFUO_ASTLO	P24913	hepesvitus
442	6	1.2	361	1	VE2_HPVT0	P50773	human papil	515	6	1.2	409	1	RNAG_HUMAN	P16634	astasia ion
443	6	1.2	361	1	DCAM_NICSY	080402	nicotiana s	516	6	1.2	410	1	AMPS_BACSV	Q02094	homo sapien
444	6	1.2	362	1	DCAM_TOBAC	004009	nicotiana t	517	6	1.2	410	1	HMH2_DROME	P33972	baicillus su
445	6	1.2	362	1	DCAM_DISTR	096555	datura stra	518	6	1.2	410	1	PGIR_MAIZE	P10035	synchocyst
446	6	1.2	363	1	Y593_CHLPN	09z7w1	chlamydia p	519	6	1.2	410	1	PGIS_MAIZE	P26216	zea mays (m
447	6	1.2	363	1	BAS8_ECOLI	P30844	escherichia	520	6	1.2	411	1	STRD_SOYBN	P33338	zea mays (m
448	6	1.2	363	1	DCAM_EPIOL	P46255	sphingia ol	521	6	1.2	412	1	AAT_RICPR	Q42807	glycine max
449	6	1.2	363	1	MURG_BACSV	P37585	baicillus su	522	6	1.2	412	1	CAR2_CANPA	Q94566	ricicetia
450	6	1.2	363	1	OMPF_SALTI	056113	salmonella	523	6	1.2	413	1	FENR_SYNY3	P33950	candida par
451	6	1.2	363	1	OMPF_SALTY	P37432	salmonella	524	6	1.2	413	1	GAT1_RAT	053138	synchocyst
452	6	1.2	366	1	SVY_SULSO	P95982	sulfolobus	525	6	1.2	416	1	CCA_HAEIN	P43429	rattus norv
453	6	1.2	369	1	HIP_HUMAN	P50802	homo sapien	526	6	1.2	417	1	IF_RAT	P45269	hemophilus
454	6	1.2	369	1	MAR2_RAT	P54404	rattus norv	527	6	1.2	417	1	KCRU_CHICK	P1267	rattus norv
455	6	1.2	369	1	SERC_HELFX	025436	heplocobact	528	6	1.2	417	1	KCRU_HUMAN	P70079	galinus gall
456	6	1.2	369	1	TMAF_AVISA	P23091	avlan muscu	529	6	1.2	418	1	KCRU_MOUSE	P12532	homo sapien
457	6	1.2	369	1	VM07_BORHE	P21876	borrelia he	530	6	1.2	418	1	KCRU_MOUSE	P30275	mus musculu
458	6	1.2	369	1	YDHH_ECOLI	P77570	escherichia	531	6	1.2	419	1	PEL_BACSV	P25809	rattus norv
459	6	1.2	370	1	EXOH_RHIME	P33692	rhizobium m	532	6	1.2	420	1	PEL_BACSV	057493	hemophilus
460	6	1.2	370	1	MAR2_MOUSE	P54843	mus musculu	533	6	1.2	421	1	PGIR_MEDSA	P33116	baicillus su
461	6	1.2	372	1	YBDK_ECOLI	P77213	galinus gall	534	6	1.2	423	1	OPPD_MYCPN	Q40312	medicago sa
462	6	1.2	373	1	TGFI_CHICK	P09531	galinus gall	535	6	1.2	423	1	SVS_COXBU	P75552	mycoplasma
463	6	1.2	374	1	FOS_TETFL	091496	tetradon f	536	6	1.2	423	1	TIG_BACSV	P33919	coxiella bu
464	6	1.2	374	1	GB15_HUMAN	P30679	homo sapien	537	6	1.2	424	1	ENO_CHLTR	P86698	baicillus su
465	6	1.2	374	1	HART_YEAST	Q12241	saccharomyc	538	6	1.2	425	1	GAZ2_MYCTU	084591	chlamydia t
466	6	1.2	375	1	YHNE_YEAST	P38797	saccharomyc	539	6	1.2	426	1	Y958_METJA	053615	mycobacteri
467	6	1.2	375	1	Y785_METJA	058195	methanococ	540	6	1.2	427	1	PRIS_METTH	Q27502	methanococ
468	6	1.2	376	1	CISX_PYRFU	053454	pyrococcus	541	6	1.2	427	1	VP8_MTV	P17380	wound tumor
469	6	1.2	376	1	FOS_FUGRU	P53450	fugu rubrip	542	6	1.2	428	1	EFIA_METJA	057770	methanococ
470	6	1.2	376	1	OMPC_SERMA	054471	serattia ma	543	6	1.2	428	1	ENO_CHLPN	Q92766	chlamydia p
471	6	1.2	377	1	CAH1_CHLRE	P20507	chlamydomon	544	6	1.2	428	1	ENO_NITEU	085348	nltrosomona

545	6	1.2	435	1	CG65_YEAST	P30283	saccharomyc	618	6	1.2	488	1	PBB_ALCEA	P12625	alcaligenes
546	6	1.2	436	1	CP22_HORVU	P55748	hordeum_vul	619	6	1.2	488	1	RBL1_CYACA	P37393	cyanidium c
547	6	1.2	436	1	MTBR_BACSU	P06530	bacillifug su	620	6	1.2	488	1	RBL_ECTSI	P2313	ectocarpus
548	6	1.2	436	1	TBD1_CORGL	P04513	corynebacte	621	6	1.2	488	1	RBL_GUTHI	P14957	guillierdia
549	6	1.2	438	1	MPK5_HUMAN	Q13163	homo sapien	622	6	1.2	488	1	RBL_OLITU	P14959	olisthodisc
550	6	1.2	442	1	TBB_TRXCR	P08562	trypanosoma	623	6	1.2	488	1	RBL_PLECA	Q08051	pleurochrys
551	6	1.2	443	1	FLI1_AQUAE	O67531	aquiflex aeo	624	6	1.2	488	1	RBL_PORAE	O09119	porphyridiu
552	6	1.2	443	1	YLM5_CAEEL	P34379	caenorhabdi	625	6	1.2	488	1	RBL_PORPU	P51226	porphyra pu
553	6	1.2	444	1	DHNA_HAEIN	P44856	haemophilus	626	6	1.2	488	1	RBL_PYLTI	P23651	pylataella p
554	6	1.2	444	1	PVDA_BURCE	O51940	burkholderi	627	6	1.2	488	1	VE2_HPV49	P36795	human papil
555	6	1.2	444	1	YNBL_YEAST	P53980	saccharomyc	628	6	1.2	489	1	FLIC_SALTY	P06179	salmonella
556	6	1.2	445	1	ACSC_WOOTH	Q07340	moorella th	629	6	1.2	489	1	PEN3_ADEMI	O10433	mouse adeno
557	6	1.2	445	1	ALCP_BACP3	P30145	bacillus ps	630	6	1.2	490	1	DHAB_BACSU	P71016	bacillus su
558	6	1.2	445	1	TBB_LEIME	P21148	leishmania	631	6	1.2	490	1	RBL_CYLSN	P4673	cyllindroche
559	6	1.2	448	1	GNTF_BACLI	P46832	bacillus li	632	6	1.2	490	1	RBL_ODOSI	P09560	mycobacteri
560	6	1.2	448	1	GNTF_BACSU	P12012	bacillus su	633	6	1.2	491	1	ACHE_BOVIN	P02715	bos taurus
561	6	1.2	449	1	OCF6_HUMAN	Q03052	homo sapien	634	6	1.2	491	1	CPH1_CHICK	P05180	gallus gall
562	6	1.2	449	1	OCF6_MOUSE	P21952	mus musculu	635	6	1.2	491	1	CPH2_CHICK	P20678	gallus gall
563	6	1.2	450	1	G6P1_BACSU	P80860	bacillus su	636	6	1.2	491	1	Y084_MYCTU	O53209	mycobacteri
564	6	1.2	450	1	VIME_CARAV	P46673	carassius a	637	6	1.2	492	1	CP53_PIG	Q02390	sus scrofa
565	6	1.2	450	1	YIE9_YEAST	P47048	saccharomyc	638	6	1.2	492	1	FLIC_SALRU	P06175	salmonella
566	6	1.2	451	1	OCF6_RAT	P20267	rattus norv	639	6	1.2	493	1	ACHE_HUMAN	O04844	homo sapien
567	6	1.2	452	1	TRPC_ECOLI	P00909	escherichia	640	6	1.2	493	1	ACHE_MOUSE	P20782	mus musculu
568	6	1.2	452	1	TRPC_SALTY	P00910	salmonella	641	6	1.2	493	1	FLIC_SALPA	P06178	salmonella
569	6	1.2	454	1	APY_SOLUT	P80595	solanum tub	642	6	1.2	494	1	ACHE_RAT	P09560	rattus norv
570	6	1.2	454	1	PR11_SCHPO	O14215	schizosacch	643	6	1.2	495	1	TOLC_ECOLI	P03293	escherichia
571	6	1.2	454	1	PUCG_RHOSH	P95656	rhodovulum	644	6	1.2	496	1	BAFI_KIDNA	P33293	kluyveromyc
572	6	1.2	455	1	VIME_CYPCA	O92155	cyprinus ca	645	6	1.2	498	1	CB75_ARATH	O65784	arabidopsis
573	6	1.2	457	1	PM4H_CAEEL	P90925	caenorhabdi	646	6	1.2	498	1	POLG_DEN19	P27909	dengue viru
574	6	1.2	458	1	DG17_DICDI	P11467	dictyosteli	647	6	1.2	499	1	MYVIN_AQUAE	O67658	aquiflex aeo
575	6	1.2	459	1	PUCG_RHOSH	Q02443	rhodobacter	648	6	1.2	500	1	FLJB_SALAE	P52615	salmonella
576	6	1.2	459	1	RBL_CALSH	P48687	calyptrosoph	649	6	1.2	500	1	PGFI_BOVIN	O29662	bos taurus
577	6	1.2	459	1	YAI2_STRCO	O86835	streptomyce	650	6	1.2	500	1	PGFI_HUMAN	O16647	homo sapien
578	6	1.2	460	1	DCED_RHIME	P13632	rhizobium m	651	6	1.2	501	1	DIDH_PEA	P33023	pisum sativ
579	6	1.2	460	1	VL2_HPV44	O80918	human papil	652	6	1.2	502	1	K2M3_SHEEP	P23691	ovis aries
580	6	1.2	460	1	VL2_HPV55	O80939	human papil	653	6	1.2	503	1	OOXA_AGR4	Q59160	agrobacteri
581	6	1.2	461	1	EFIA_ARTSA	P02993	artemisa lae	654	6	1.2	504	1	FLIC_SALMU	P06177	salmonella
582	6	1.2	461	1	PR58_XENLA	P46770	xenopus lae	655	6	1.2	504	1	SAH2_DROME	P50245	dirosophila
583	6	1.2	462	1	TUBE_DROME	P22812	dirosophila	656	6	1.2	504	1	TLR1_DROME	P30974	dirosophila
584	6	1.2	465	1	EXSH_RHIME	O33680	rhizobium m	657	6	1.2	505	1	C762_SOLME	P37122	solanum mel
585	6	1.2	465	1	SYN_ECOLI	P17242	escherichia	658	6	1.2	505	1	FLJB_SALTY	P52616	salmonella
586	6	1.2	465	1	YB57_YEAST	P38110	saccharomyc	659	6	1.2	509	1	YRM4_CAEEL	Q10051	caenorhabdi
587	6	1.2	466	1	CFIA_DROME	P16241	dirosophila	660	6	1.2	510	1	FKH_DROME	P14734	dirosophila
588	6	1.2	467	1	CISY_CANTR	P79924	candida tro	661	6	1.2	512	1	DDDH_SCHPO	O00087	schizosacch
589	6	1.2	468	1	GLNA_AZOB	P10583	azospirillu	662	6	1.2	512	1	K2C5_XENLA	P16878	xenopus lae
590	6	1.2	468	1	GLNA_ECOLI	P06611	escherichia	663	6	1.2	512	1	VP67_NPVAC	P17501	autographa
591	6	1.2	468	1	GLNA_SALTY	P06601	salmonella	664	6	1.2	513	1	TYRR_SALTY	Q05427	schizosacch
592	6	1.2	468	1	GLNA_THIFE	P07804	thiobacillu	665	6	1.2	513	1	Y073_TREPA	O83112	treponema p
593	6	1.2	469	1	GLNA_PROVU	P28786	proteus vul	666	6	1.2	514	1	VS14_TRYBB	P28329	trypanosoma
594	6	1.2	469	1	PPA5_KLUFA	P52289	kluyveromyc	667	6	1.2	514	1	VYJI_HAEIN	P44744	haemophilus
595	6	1.2	470	1	YMR7_YEAST	O04371	saccharomyc	668	6	1.2	517	1	CP53_ASPNG	P17549	aspergillus
596	6	1.2	472	1	GLNA_HAEIN	P43794	haemophilus	669	6	1.2	517	1	EGG1_YEAST	P32474	saccharomyc
597	6	1.2	472	1	IFR2_HUMAN	P09913	homo sapien	670	6	1.2	517	1	Y073_TREPA	O83112	treponema p
598	6	1.2	473	1	LCB1_CRIGR	O54695	cricketul	671	6	1.2	518	1	ATPB_NEUCR	P23704	neurospora
599	6	1.2	473	1	LCB1_HUMAN	O15659	homo sapien	672	6	1.2	519	1	YX23_CAEEL	P14734	dirosophila
600	6	1.2	473	1	LCB1_MOUSE	O35704	mus musculu	673	6	1.2	519	1	CP5K_CANNA	Q12589	candida mal
601	6	1.2	474	1	YPC2_CAEEL	P15531	mus musculu	674	6	1.2	520	1	TRM1_ECOLI	P10484	escherichia
602	6	1.2	475	1	PERI_MOUSE	P53001	saccharomyc	675	6	1.2	520	1	TRM1_FUSSP	Q16612	fusarium sp
603	6	1.2	476	1	YGI2_YEAST	P42305	bacillus su	676	6	1.2	520	1	YVAV_CAEEL	O27517	caenorhabdi
604	6	1.2	479	1	DEAD_BACSU	P77416	escherichia	677	6	1.2	522	1	ABP1_SCHPO	P49777	schizosacch
605	6	1.2	480	1	HYFD_ECOLI	P11172	homo sapien	678	6	1.2	522	1	CK13_YEAST	P39962	saccharomyc
606	6	1.2	480	1	PYR5_HUMAN	O00217	aspergillus	679	6	1.2	524	1	CP5F_CANTR	P36068	candida tro
607	6	1.2	482	1	TPSB_ASPNG	P11181	bos taurus	680	6	1.2	524	1	GLPK_HUMAN	P35189	homo sapien
608	6	1.2	482	1	ODB2_BOVIN	P11182	homo sapien	681	6	1.2	524	1	GLPK_MOUSE	O63516	mus musculu
609	6	1.2	482	1	ODB2_HUMAN	P53395	mus musculu	682	6	1.2	524	1	DPE2_CAEEL	O19196	caenorhabdi
610	6	1.2	482	1	YCI1_KLEPN	O48457	klebsiella	683	6	1.2	526	1	MS51_YEAST	P33559	saccharomyc
611	6	1.2	485	1	RBL1_RHOSH	P27997	rhodospacter	684	6	1.2	526	1	OOXA_RHIME	P72300	rhizobium m
612	6	1.2	486	1	RBL1_RHOSH	P09657	alcaligenes	685	6	1.2	526	1	NGGH_BACSU	P42954	bacillus su
613	6	1.2	486	1	RBL1_ALCEU	P42721	alcaligenes	686	6	1.2	527	1	TF65_XENLA	O04865	xenopus lae
614	6	1.2	486	1	RBL_RHIME	P56689	rhizobium m	687	6	1.2	527	1	NIFK_FRALV	O57118	frankia aln
615	6	1.2	486	1	RBL_RHIME	P47195	berberis st	688	6	1.2	528	1	UDB9_MACFA	O02663	macaca fasc
616	6	1.2	487	1	CP80_BERST	P06176	salmonella	689	6	1.2	529	1			
617	6	1.2	488	1	FLIC_SALCH			690	6	1.2	529	1			

691	6	1.2	530	1	UDBC_RAT	P36511	rattus norv	764	6	1.2	630	1	YD13_SCHPO	Q92341	schizosacch
692	6	1.2	530	1	UDBF_HUMAN	P34855	homo sapien	765	6	1.2	633	1	DX5_BACSU	P54523	bacillus su
693	6	1.2	530	1	UDBF_HUMAN	O75795	homo sapien	766	6	1.2	635	1	KSYK_HUMAN	P43305	homo sapien
694	6	1.2	531	1	CP27_HUMAN	O02318	h cytochrom	767	6	1.2	636	1	PHBC_RIET	O52728	r poly-beta
695	6	1.2	535	1	UD11_MOUSE	O63886	mus musculu	768	6	1.2	638	1	GHR_HUMAN	P10912	homo sapien
696	6	1.2	538	1	IPGD_MOUSE	O07566	shigella fl	769	6	1.2	638	1	GHR_MACMU	P79194	macca mula
697	6	1.2	538	1	IPGD_SHISO	O55286	shigella so	770	6	1.2	638	1	KRAF_XENLA	P09560	xenopus lae
698	6	1.2	538	1	YABK_HAETI	P44963	haemophilus	771	6	1.2	640	1	YIF6_YEAST	P40522	saccharomyc
699	6	1.2	540	1	NADB_ECOLI	P10902	escherichia	772	6	1.2	643	1	GYRE_MYCA	P50028	mycoplasma
700	6	1.2	542	1	THS_METJA	O58405	methanococc	773	6	1.2	643	1	RHOE_MOUSE	O61825	mus musculu
701	6	1.2	547	1	ATP2_DAUCA	P73399	daucus caro	774	6	1.2	644	1	FTSH_ODOSI	P49825	odontella s
702	6	1.2	547	1	BUDP_YEAST	P53236	saccharomyc	775	6	1.2	644	1	RNR_HELPE	P56123	helicobacte
703	6	1.2	547	1	CH60_PSEAE	P30718	pseudomonas	776	6	1.2	648	1	AMPB_RAT	O09173	rattus norv
704	6	1.2	549	1	TF65_MOUSE	O04207	mus musculu	777	6	1.2	648	1	DPOL_BPSF2	P06225	bacterioph
705	6	1.2	549	1	YST1_CAEEL	O22000	caenorhabdi	778	6	1.2	648	1	MUTL_TREPA	O83325	treponea p
706	6	1.2	550	1	YCHM_ECOLI	P40877	escherichia	779	6	1.2	652	1	DIYN_DICDI	P54303	dicyostell
707	6	1.2	551	1	TF65_HUMAN	O04206	homo sapien	780	6	1.2	653	1	LEPA_MYCTU	P71739	mycobacteri
708	6	1.2	555	1	NEL_CORTJA	O02916	coturnix co	781	6	1.2	668	1	COAT_FCVF6	P27405	feline call
709	6	1.2	555	1	POLG_DENIT	P29983	dengue viru	782	6	1.2	668	1	COAT_FCVF4	P27405	feline call
710	6	1.2	556	1	PLD_STRAT	O53728	streptomyce	783	6	1.2	669	1	ILVB_SCHPO	P36620	schizosacch
711	6	1.2	556	1	RSI_HELPE	P56008	helicobacte	784	6	1.2	670	1	KGPA_BOVIN	O13976	homo sapien
712	6	1.2	556	1	SYDC_YEAST	O4802	saccharomyc	785	6	1.2	670	1	KGPA_HUMAN	O13976	homo sapien
713	6	1.2	557	1	MK04_HUMAN	P31152	homo sapien	786	6	1.2	670	1	KGPA_RABIT	O76756	oryctolagus
714	6	1.2	559	1	ENL_HUMAN	O03111	homo sapien	787	6	1.2	671	1	COAT_FCVF9	P27406	feline call
715	6	1.2	561	1	ILVD_METJA	O58672	methanococc	788	6	1.2	673	1	SIM_DROME	P05709	drosophila
716	6	1.2	561	1	SYR_CHLEP	O92773	chlamydia p	789	6	1.2	675	1	HS7M_PEA	P37300	pisum sativ
717	6	1.2	563	1	P0UT2_HUMAN	P30038	homo sapien	790	6	1.2	675	1	HS7M_PHAUV	Q01899	phaseolus v
718	6	1.2	565	1	HEMA_IAGUA	P31011	influenza a	791	6	1.2	676	1	CCME_RHIME	P45404	rhizobium m
719	6	1.2	566	1	AMY_STRGR	P30270	streptomyce	792	6	1.2	676	1	IF2M_YEAST	P25038	saccharomyc
720	6	1.2	566	1	AMY_STRLM	P09794	streptomyce	793	6	1.2	676	1	TIM_DROHY	O44431	drosophila
721	6	1.2	566	1	BGIC_MAIZE	P49235	zea mays (m	794	6	1.2	677	1	BGAL_HUMAN	P16378	homo sapien
722	6	1.2	566	1	HEMA_IAGU2	P31303	influenza a	795	6	1.2	678	1	YRAN_ECOLI	P43564	escherichia
723	6	1.2	566	1	HEMA_IAPIL	P13102	influenza a	796	6	1.2	678	1	TKTA_CRAPL	O42675	craterostig
724	6	1.2	572	1	HLV1_AERSA	O08675	aeromonas s	797	6	1.2	682	1	HS7M_SOLTU	O08976	solanum tub
725	6	1.2	573	1	CH61_DROME	O02649	drosophila	798	6	1.2	683	1	RPC_BPPHC	P08979	bacterioph
726	6	1.2	573	1	P60_CRIGR	P18687	cricketulus	799	6	1.2	686	1	KGPA_BOVIN	P21136	bos taurus
727	6	1.2	573	1	P60_MOUSE	P19226	mus musculu	800	6	1.2	686	1	KGPA_HUMAN	P14619	homo sapien
728	6	1.2	574	1	MAOI_ECOLI	P26616	escherichia	801	6	1.2	686	1	KGPA_MOUSE	Q92020	mus musculu
729	6	1.2	576	1	YAG3_SCHPO	O09868	schizosacch	802	6	1.2	687	1	YB80_YEAST	P34325	saccharomyc
730	6	1.2	579	1	SYD_HELPE	O92119	helicobacte	803	6	1.2	688	1	YJ80_YEAST	P47140	saccharomyc
731	6	1.2	583	1	YCV1_YEAST	P25639	saccharomyc	804	6	1.2	700	1	LCFL_YEAST	P30624	saccharomyc
732	6	1.2	585	1	CH60_PYRSA	P46224	pyrenomonas	805	6	1.2	701	1	CGI_HUMAN	O13495	homo sapien
733	6	1.2	586	1	ACES_TORCA	P04058	torpedo cal	806	6	1.2	701	1	HS83_LETAM	P27741	leishmania
734	6	1.2	586	1	SECD_BORBU	O51536	borrella bu	807	6	1.2	702	1	COAT_SMSV1	P36284	san miguel
735	6	1.2	586	1	Y187_MYCPN	P75264	mycoplasma	808	6	1.2	703	1	COAT_SMSV4	P36284	san miguel
736	6	1.2	590	1	ACES_TORMA	P07692	torpedo mar	809	6	1.2	710	1	ETEF2_VACCC	P20635	vaccinia vl
737	6	1.2	590	1	PEPF_BORBU	O51264	borrella bu	810	6	1.2	710	1	ETEF2_VARY	P33806	variola vir
738	6	1.2	592	1	EVAL_HUMAN	O99502	homo sapien	811	6	1.2	712	1	RNR_THEMA	O92411	thermotoga
739	6	1.2	596	1	DHSA_RICPR	P31038	rickettsia	812	6	1.2	712	1	DNL2_RHOMR	P49421	rhodothermu
740	6	1.2	596	1	HMEN_ANGA	O02451	anopheles g	813	6	1.2	719	1	MUS2_STRAU	Q92455	staphylococ
741	6	1.2	600	1	DHSA_PANDE	O59661	paracoccuss	814	6	1.2	719	1	PBPA_STRPN	O47407	streptococc
742	6	1.2	601	1	PDML_DROME	P31368	drosophila	815	6	1.2	720	1	SPOT_MYGGE	P47520	mycoplasma
743	6	1.2	601	1	T2S1_STRSA	P29346	streptococc	816	6	1.2	720	1	YMS2_YEAST	O10767	saccharomyc
744	6	1.2	604	1	AMVG_RHIOA	P07683	rhizopus or	817	6	1.2	721	1	GLGX_MYCTU	O10767	saccharomyc
745	6	1.2	604	1	NODM_RHIME	P25135	rhizobium m	818	6	1.2	722	1	ACML_DROME	P16395	drosophila
746	6	1.2	604	1	YFIC_BACSU	P54719	bacillus su	819	6	1.2	732	1	COPA_HELFE	O32619	helicobacte
747	6	1.2	606	1	HMDI_DROAN	P22514	drosophila	820	6	1.2	733	1	MUTB_STRCM	Q05065	streptomyce
748	6	1.2	609	1	KMLC_RAT	P20689	rattus norv	821	6	1.2	733	1	YACK_RHIME	O92442	rhizobium m
749	6	1.2	610	1	ARLC_MATZE	P13536	zea mays (m	822	6	1.2	735	1	FCF_ERMCH	O47162	erythra chr
750	6	1.2	612	1	ARRS_MAIZE	O53250	maizeobacteri	823	6	1.2	741	1	CH12_YEAST	P49956	saccharomyc
751	6	1.2	618	1	ILVB_MYCTU	O02335	caenorhabdi	824	6	1.2	741	1	COA1_HELPY	O59467	helicobacte
752	6	1.2	619	1	YOLB_CAEEL	P23543	homo sapien	825	6	1.2	745	1	COA1_HELPY	P77871	helicobacte
753	6	1.2	619	1	ZAT0_HUMAN	P28604	a nodg bifu	826	6	1.2	745	1	COA3_HELPY	O08462	helicobacte
754	6	1.2	620	1	NODO_AZOB	O59498	mycobacteri	827	6	1.2	748	1	CAO_YEAST	P13711	saccharomyc
755	6	1.2	621	1	ILVB_MYCAV	O03447	mycobacteri	828	6	1.2	750	1	CTPB_MYCLE	P46640	mycobacteri
756	6	1.2	621	1	LAMC_DROME	O03447	drosophila	829	6	1.2	757	1	ECR_LUTCU	O18531	lucilia cup
757	6	1.2	622	1	LAMC_DROME	P08928	drosophila	830	6	1.2	757	1	MUS2_THEMA	Q924105	thermotoga
758	6	1.2	622	1	VP27_YEAST	P45033	saccharomyc	831	6	1.2	758	1	HUNB_DROME	P08084	drosophila
759	6	1.2	625	1	BAR3_SCHCO	P56502	schizosacch	832	6	1.2	761	1	METE_BACSU	O61337	homo sapien
760	6	1.2	626	1	PPOC_LYCES	O08305	lycopersico	833	6	1.2	762	1	KGPA_HUMAN	O61337	homo sapien
761	6	1.2	628	1	KSYK_PIG	O00655	sus scrofa	834	6	1.2	762	1	KGPA_MOUSE	O64595	mus musculu
762	6	1.2	630	1	GIMS_SVNY3	P72720	s glucosam1	835	6	1.2	762	1	KGPA_RAT	O64595	rattus norv
763	6	1.2	630	1	PPOR_LYCES	O08303	lycopersico	836	6	1.2	763	1	IF2C_PORPU	P51257	porphyra pu

837	6	1.2	764	1	DRA_HUMAN	P40879	homo sapien	910	6	1.2	919	1	VGJB_HSV4	P17472	equine herp
838	6	1.2	770	1	STAG_HUMAN	P40763	homo sapien	911	6	1.2	924	1	ORC1_DROME	O16810	drosophila
839	6	1.2	770	1	STAG_MOUSE	P42227	mus musculus	912	6	1.2	928	1	KINH_NEUCR	P48467	neuropora
840	6	1.2	770	1	STAG_RAT	P52631	rattus norv	913	6	1.2	931	1	CA19_HUMAN	P20849	homo sapien
841	6	1.2	771	1	YJCO_YEAST	P47068	saccharomyc	914	6	1.2	933	1	PERT_HUMAN	P07202	homo sapien
842	6	1.2	773	1	GYRB_HELPJ	Q921X3	helicobacte	915	6	1.2	933	1	VGJB_HSV1	O04463	herpesvirus
843	6	1.2	773	1	GYRB_HELPJ	P55992	helicobacte	916	6	1.2	934	1	CITC_HUMAN	P11586	h-c-1-tetra
844	6	1.2	774	1	SCA_DROME	P21520	drosophila	917	6	1.2	936	1	CITC_RAT	Q94653	ricchetti
845	6	1.2	776	1	GCR_XENLA	P49844	xenopus lae	918	6	1.2	934	1	ODOL_RICPR	Q94653	ricchetti
846	6	1.2	776	1	VP4_ROTBU	P12474	bovine rota	919	6	1.2	938	1	YV91_CAEEL	O22712	caenorhabd
847	6	1.2	776	1	VP4_ROTBU	P12474	bovine rota	920	6	1.2	942	1	PR11_HUMAN	O22712	caenorhabd
848	6	1.2	779	1	RNR_BACSU	O08778	human rotav	921	6	1.2	943	1	ODOL_AZOV1	O15512	homo sapien
849	6	1.2	780	1	TREB_YEAST	P35172	saccharomyc	922	6	1.2	952	1	UVRA_THER1	P20707	azotobacter
850	6	1.2	780	1	RIRI_HELPJ	Q921F9	helicobacte	923	6	1.2	953	1	YV91_YEAST	O56242	thermus agu
851	6	1.2	788	1	RIRI_HELPJ	P55982	helicobacte	924	6	1.2	957	1	NIRB_KLEPN	P07202	homo sapien
852	6	1.2	788	1	TRPG_PCHAC	P25170	p anthranil	925	6	1.2	963	1	KINH_HUMAN	O04658	kibsiella
853	6	1.2	791	1	POLG_DENIC	P27913	dengue viru	926	6	1.2	963	1	KINH_HUMAN	P33176	homo sapien
854	6	1.2	792	1	POLG_DEN18	P27910	dengue viru	927	6	1.2	964	1	DPOL_CBEPY	Q61768	mus musculu
855	6	1.2	792	1	POLG_DEN18	P27910	dengue viru	928	6	1.2	967	1	SOXA_COR51	P30319	chortistoneu
856	6	1.2	797	1	CPE_MYCTU	O08365	mycobacteri	929	6	1.2	969	1	SACB_STRSL	Q55242	streptococc
857	6	1.2	797	1	DPOM_AGABT	P30322	agarcus bl	930	6	1.2	975	1	CITM_YEAST	P08440	s-c-1-tetra
858	6	1.2	797	1	VGIX_HSVB	P28968	equine herp	931	6	1.2	979	1	GLK5_MOUSE	O61626	mus musculu
859	6	1.2	798	1	HEBA_HOMVA	P16827	human cytom	932	6	1.2	979	1	GLK5_RAT	O61626	mus musculu
860	6	1.2	799	1	HIS2_YEAST	P00815	saccharomyc	933	6	1.2	979	1	VGJB_HSV4	P18551	equine herp
861	6	1.2	801	1	PIPA_DICDI	Q02158	dictyostell	934	6	1.2	980	1	GLK5_HUMAN	P23218	equine herp
862	6	1.2	801	1	GYRB_CHLTR	O84193	chlamydia t	935	6	1.2	980	1	VGJB_HSV1	P18050	equine herp
863	6	1.2	805	1	PR1A_BACSU	P94461	bacillus su	936	6	1.2	980	1	VGJB_HSV4	P18551	equine herp
864	6	1.2	812	1	FGRI_XENLA	P22182	xenopus lae	937	6	1.2	980	1	VGJB_HSV4	P18551	equine herp
865	6	1.2	813	1	FPVA_PSEAE	P48632	pseudomonas	938	6	1.2	982	1	POL_FITV2	P23922	equine herp
866	6	1.2	816	1	HUMB_DROVI	P13361	drosophila	939	6	1.2	985	1	AGLU_ASPNG	P03363	human t-cel
867	6	1.2	816	1	QALF_NEUCR	P11638	neuropora	940	6	1.2	985	1	EPB2_HUMAN	P56526	aspergillus
868	6	1.2	819	1	FGRI_CHICK	P21804	gallus gall	941	6	1.2	986	1	EP1B_STAPB	P29323	homo sapien
869	6	1.2	820	1	CH1A_ALTSO	P32823	alteromonas	942	6	1.2	987	1	EPB2_COTJA	O90045	staphylococ
870	6	1.2	820	1	FES_MOUSE	P14338	felis silve	943	6	1.2	987	1	YD94_METVA	O58789	methanococ
871	6	1.2	820	1	FES_MOUSE	P16879	mus musculu	944	6	1.2	988	1	EPB2_CHICK	P28693	gallus gall
872	6	1.2	821	1	YQIG_ECOLI	P76655	escherichia	945	6	1.2	989	1	SMB2_MESAU	O60560	mesocricetu
873	6	1.2	822	1	FGRI_MOUSE	P11662	homo sapien	946	6	1.2	993	1	AFER_STRCO	P25941	streptomyce
874	6	1.2	822	1	FGRI_MOUSE	P16092	mus musculu	947	6	1.2	993	1	EPB2_MOUSE	P54763	mus musculu
875	6	1.2	822	1	FGRI_MOUSE	O04589	rattus norv	948	6	1.2	993	1	SMB2_MOUSE	P40694	mus musculu
876	6	1.2	822	1	Y539_HUMAN	O60287	homo sapien	949	6	1.2	997	1	VGMA_APMV	P38485	andean pola
877	6	1.2	830	1	CC19_SCHPO	P40377	schizosacch	950	6	1.2	999	1	ATC3_HUMAN	O93084	homo sapien
878	6	1.2	830	1	MKT1_YEAST	P40850	saccharomyc	951	6	1.2	1002	1	ATCB_DROME	P18596	rattus norv
879	6	1.2	831	1	NAPA_RHOSH	O53176	rhodobacter	952	6	1.2	1002	1	ATC3_RAT	P22700	drosophila
880	6	1.2	831	1	YPDD_ECOLI	P77439	escherichia	953	6	1.2	1003	1	ATC3_RAT	P3316	artemia san
881	6	1.2	832	1	TRNL_CANAL	P43075	candida alb	954	6	1.2	1003	1	SYG_CHLTR	O46371	chlamydia t
882	6	1.2	835	1	Y422_MYCE	P47661	mycoplasma	955	6	1.2	1004	1	PO11_SCTCO	O03377	scitara copr
883	6	1.2	844	1	APB2_YEAST	P32454	saccharomyc	956	6	1.2	1016	1	UVRA_DEIRA	O46577	deinococcus
884	6	1.2	845	1	SCPI_MESAU	O60563	mesocricetu	957	6	1.2	1021	1	Y2R2_DROME	P16425	drosophila
885	6	1.2	851	1	ENV_HV1B8	P04582	human immu	958	6	1.2	1030	1	FBP1_DROME	O04691	drosophila
886	6	1.2	856	1	ENV_HV1B2	P04578	human immu	959	6	1.2	1032	1	Y0S3_CAEEL	O09249	caenorhabd
887	6	1.2	856	1	ENV_HV1B3	P04624	human immu	960	6	1.2	1035	1	POLY_DROME	P10401	drosophila
888	6	1.2	859	1	ABR_HUMAN	Q12379	homo sapien	961	6	1.2	1036	1	Y414_MYCE	P47653	mycoplasma
889	6	1.2	861	1	GCR3_YEAST	P34160	saccharomyc	962	6	1.2	1039	1	MSL1_DROME	P50535	drosophila
890	6	1.2	872	1	VP2_ROTBC	P26191	porcine rot	963	6	1.2	1046	1	CH1D_VIBFU	P96156	virio furn
891	6	1.2	874	1	SLAP_BACLI	P49052	bacillus il	964	6	1.2	1048	1	P100_HOMVA	P00818	human cytom
892	6	1.2	875	1	H1R2_YEAST	P32480	saccharomyc	965	6	1.2	1053	1	CARB_SCHPO	O10283	schizosacch
893	6	1.2	876	1	SYA_ECOLI	P00957	escherichia	966	6	1.2	1058	1	CARB_LACP1	P77886	lactobacill
894	6	1.2	877	1	DPOL_LACIC	O32801	lactococcus	967	6	1.2	1060	1	EG51_XENLA	P28025	xenopus lae
895	6	1.2	877	1	INCE_CHICK	P33532	gallus gall	968	6	1.2	1067	1	EG52_XENLA	O91783	xenopus lae
896	6	1.2	877	1	SYA_RICPR	O59637	pseudomonas	969	6	1.2	1072	1	MAP4_BOVIN	P36525	bos taurus
897	6	1.2	880	1	ODPI_PSEAE	P28277	herpes simp	970	6	1.2	1081	1	GALY_YEAST	P19659	saccharomyc
898	6	1.2	881	1	HELI_HSV2H	P28277	herpes simp	971	6	1.2	1081	1	SVY_TETTH	P36525	saccharomyc
899	6	1.2	882	1	CADI_HUMAN	P12830	homo sapien	972	6	1.2	1083	1	T2D3_HUMAN	O00268	homo sapien
900	6	1.2	884	1	CADI_MOUSE	P09803	mus musculu	973	6	1.2	1103	1	CYED_HUMAN	O00268	homo sapien
901	6	1.2	892	1	RA16_SCHPO	P36617	schizosacch	974	6	1.2	1104	1	NIR_EMENT	P22944	emeritella
902	6	1.2	892	1	YLB3_CAEEL	P46578	caenorhabd	975	6	1.2	1109	1	CYGD_CANFA	O19179	canis fami
903	6	1.2	894	1	YH1H_ECOLI	P37624	escherichia	976	6	1.2	1113	1	IMB4_YEAST	P40069	saccharomyc
904	6	1.2	896	1	CB1_MOUSE	P22682	mus musculu	977	6	1.2	1119	1	AL53_CANAL	O74625	candida alb
905	6	1.2	899	1	PR06_YEAST	P19735	saccharomyc	978	6	1.2	1120	1	AL53_CANAL	P77338	escherichia
906	6	1.2	899	1	YABD_SCHPO	O09778	schizosacch	979	6	1.2	1124	1	POL_FITPE	P16088	feline immu
907	6	1.2	904	1	SYA_MCTU	O07438	mycobacteri	980	6	1.2	1124	1	POL_FITVD	P19028	feline immu
908	6	1.2	909	1	CNG4_HUMAN	Q14028	homo sapien	981	6	1.2	1124	1	POL_FITV2	P31822	feline immu
909	6	1.2	918	1	QAL5_NEUCR	P11637	neuropora	982	6	1.2	1129	1	RPA2_DROME	P20028	drosophila

```

883 6 1.2 1132 1 NTU1_YEAST P53114 saccharomyc
984 6 1.2 1135 1 RB12_PAT 055061 rattus norv
985 6 1.2 1136 1 ADD9_BACSU P23477 bacillus su
986 6 1.2 1157 1 AGE1_CAEEL 094125 caenorhabd1
987 6 1.2 1184 1 POL2_GPELV P18474 grapevine f
988 6 1.2 1194 1 DPOL_VZVD P09252 varicella-2
989 6 1.2 1195 1 METH_SYNY3 055786 synochocyt
990 6 1.2 1199 1 N121_PAT P52501 rattus norv
991 6 1.2 1209 1 DNBI_HSYEB P28932 equine herp
992 6 1.2 1209 1 THR_DROME P42286 drosophila
993 6 1.2 1220 1 DPOL_HSYEB P28858 equine herp
994 6 1.2 1226 1 POLG_DENIW P17763 dengue viru
995 6 1.2 1248 1 TOPG_SULAC 008582 sulfolobus
996 6 1.2 1250 1 VFAL_ECOLI P45508 escherichia
997 6 1.2 1251 1 RBP2_PLAVB 000799 plasmodium
998 6 1.2 1267 1 VL3_REOVD P17378 reovirus (t
999 6 1.2 1267 1 VL3_REOVL P17376 reovirus (t
1000 6 1.2 1273 1 MYS3_YEAST P36006 saccharomyc

```

ALIGNMENTS

```

RESULT 1
LCRV_YERPE STANDARD: PRT: 326 AA.
ID LCRV_YERPE
AC P21206;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN
V).
GN LCRV.
OS Yersinia pestis.
OG Plasmid pCD1.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM:
RC MEDLINE: 90008806.
RA Price S.B., Leung K.Y., Barve S.S., Straley S.C.;
RT "Molecular analysis of lcrGVH, the V antigen operon of Yersinia
RT pestis.";
RL J. Bacteriol. 171:5646-5653(1989).
CC -1- FUNCTION: POSSIBLY INVOLVED IN CA(2+) REGULATION OF YOP
CC EXPRESSION, WHICH INCLUDES THE EXPORT PROCESS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC EMBL: M26405; AAA27641.1; ALT_SEQ.
CC DR PIR: B33601; B33601.
CC KW Plasmid; Antigen; Virulence.
CC SO SEQUENCE 326 AA; 37226 MW; 54FB8209E032F3F4 CRC64;

```

```

Query Match 36.7%; Score 191; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 1e-175;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 331 LKIVYSNMHNGARSKLRELAELTRELKITYTQAEIKHLSGGTINIHKSNIMD 330
DB 136 LKIVYSNMHNGARSKLRELAELTRELKITYTQAEIKHLSGGTINIHKSNIMD 195
QY 391 KNLGYTDEEIFRASAAYKLEKMPOTTIOVDGSEKKIVSINKPLGSENKRTGALGNKN 450
DB 136 KNLGYTDEEIFRASAAYKLEKMPOTTIOVDGSEKKIVSINKPLGSENKRTGALGNKN 450

```

```

DB 196 KNLGYTDEEIFRASAAYKLEKMPOTTIOVDGSEKKIVSINKPLGSENKRTGALGNKN 255
QY 451 SYSYNKDNNELSHFATCSDKSRPLNDVYSQKTTQSLDITSRPNISAIEALNRFIOKYSV 510
DB 256 SYSYNKDNNELSHFATCSDKSRPLNDVYSQKTTQSLDITSRPNISAIEALNRFIOKYSV 315
QY 511 MORLDDTSGK 521
DB 316 MORLDDTSGK 326

```

```

RESULT 2
CAP1_YERPE STANDARD: PRT: 170 AA.
ID CAP1_YERPE
AC P26948;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE P1 CAPSULE ANTIGEN PRECURSOR.
GN CAP1.
OS Yersinia pestis.
OG Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE: 91099503.
RA Galinov E.E., Smirnov O.Y., Karlishhev A.V., Volkovoy K.I.,
RA Denezuyuk A.I., Nazimov I.V., Rubtsov K.S., Abramov V.M.,
RA Dalavdanz S.M., Zay'yalov V.P.;
RT "Nucleotide sequence of the Yersinia pestis gene encoding P1 antigen
RT and the primary structure of the protein. Putative T and B cell
RT epitopes.";
RL FEBS Lett. 277:230-232(1990).
CC -1- SUBCELLULAR LOCATION: CAPSULE.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC EMBL: X61996; CAA43966.1; -
CC DR PIR: S13008; S13008.
CC KW Plasmid; Signal.
CC FT SIGNAL 1 21
CC FT CHAIN 22 170 POTENTIAL.
CC FT DOMAIN 100 150 CONTAINS POTENTIAL ANTIGENIC DETERMINANTS
CC FT SEQUENCE 170 AA; 17666 MW; 9AC87796A0BA67D9 CRC64;

```

```

Query Match 32.6%; Score 170; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 8.6e-156;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 24 MKRISSVIAIALFGTATNAADLTASTATATVLPARTTLTKRGAPTTINDNNDIT 83
DB 1 MKRISSVIAIALFGTATNAADLTASTATATVLPARTTLTKRGAPTTINDNNDIT 60
QY 84 ELIVGTLTIGGYTGTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRFDIS 143
DB 61 ELIVGTLTIGGYTGTSTSVNFTDAAGDPMYLTFTSODGNHQTFTKYIGKDSRFDIS 120
QY 144 PKVNGENLVGDVYVLAATGSDFFVRSIGSKGKLAAGKTTDAVTVVSNQ 193
DB 121 PKVNGENLVGDVYVLAATGSDFFVRSIGSKGKLAAGKTTDAVTVVSNQ 170

```

```

RESULT 3
LCRV_YERPE

```



```

ID  LCRV_YERPS      STANDARD:      PRT:      326 AA.
AC  P23994;
DT  01-MAR-1992 (Rel. 21, Created)
DT  01-MAR-1992 (Rel. 21, Last sequence update)
DT  15-FEB-2000 (Rel. 39, Last annotation update)
DE  VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN
DE  V).
GN  LCRV.
OS  Yersinia pseudotuberculosis.
OG  Plasmid pIB1.
OC  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC  Yersinia.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-YPIII;
RX  MEDLINE: 91154114.
RA  Bergman T., Haakansson S., Forsberg A., Norlander L., Macellaro A.,
RA  Beekman A., Boelijn I., Wolf-Watz H.
RT  "Analysis of the V antigen lcrGVH-yopBD operon of Yersinia
RT  pseudotuberculosis: evidence for a regulatory role of lcrV and
RT  lcrV."
RL  J. Bacteriol. 173:1607-1616(1991).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE: 97162308.
RA  Roggenkamp A., Geiger A.M., Leitritz L., Kessler A., Heesemann J.;
RT  "Passive immunity to infection with Yersinia spp. mediated by anti-
RT  recombinant V antigen is dependent on polymorphism of V antigen."
RL  Infect. Immun. 65:446-451(1997).
CC  -1- FUNCTION: INVOLVED IN CA(2+) REGULATION OF YOP EXPRESSION, WHICH
CC  INCLUDES THE EXPORT PROCESS.
CC  -1- SUBCELLULAR LOCATION: SECRETED.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; M57893; AAA27645.1; -
DR  EMBL; X96802; CAA65594.1; -
DR  PIR; B37314; B37314.
KW  Plasmid; Antigen; Virulence.
SQ  SEQUENCE 326 AA; 37336 MW; 2FD945DAFF48C06 CRC64;

Query Match      15.7%; Score 82; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 5.3e-71;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  196 MRAAEONPHATEDELEKRVQQLTGHGSSVLEBYQVLYKXNDISITVPRKSEVFA 255
DB  1 MRAAEONPHATEDELEKRVQQLTGHGSSVLEBYQVLYKXNDISITVPRKSEVFA 60
OY  256 NRVITDDIELKKILAYFLPED 277
DB  61 NRVITDDIELKKILAYFLPED 82

RESULT 4
MOX2_HUMAN      STANDARD:      PRT:      303 AA.
AC  P50222;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  15-JUL-1998 (Rel. 36, Last annotation update)
DE  HOMEBOX PROTEIN MOX-2 (GROWTH ARREST-SPECIFIC HOMEBOX).
GN  MOX2 OR MOX2 OR GAX.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE-EMBRYO;
RX  MEDLINE: 95331791.
RA  Grigoriou M., Kastirnakl M.-C., Modl W., Theodorakis K., Mankoo B.,
RA  Pachnis V., Karagogeos D.;
RT  "Isolation of the human MOX2 homeobox gene and localization to
RT  chromosome 7p22.1-p21.3."
RL  Genomics 26:550-555(1995).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE-HEART;
RX  MEDLINE: 95229154.
RA  Lepage D.F., Walsh K.;
RT  "Molecular cloning and localization of the human GAX gene to 7p21."
RL  Genomics 24:535-540(1994).
CC  -1- FUNCTION: ROLE IN MESODERM INDUCTION AND ITS EARLIEST REGIONAL
CC  SPECIFICATION, SOMITOGENESIS, AND MYOGENIC AND SCLEROTOMAL
CC  DIFFERENTIATION. MAY HAVE A REGULATORY ROLE WHEN QUIESCENT
CC  VASCULAR SMOOTH MUSCLE CELLS REENTER THE CELL CYCLE (BY
CC  SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC  -1- TISSUE SPECIFICITY: EMBRYO AND PLACENTA.
CC  -1- SIMILARITY: WITH OTHER HOMEBOX PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; X82629; CAA57949.1; -
DR  EMBL; L36328; AAB58497.1; -
DR  HSSP; P02833; ISAN.
DR  MIM; 600535; -
DR  PFM; PF00046; homeobox.1.
DR  PRINTS; PR00024; HOMEBOX.
DR  PROSITE; PS00027; HOMEBOX.1; 1.
DR  PROSITE; PS50071; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Nuclear protein; Developmental protein.
FT  DOMAIN 42 47 POLY-SER.
FT  DOMAIN 68 79 POLY-HIS.
FT  DOMAIN 80 85 POLY-GLN.
FT  DNA_BIND 186 245 HOMEBOX.
FT  CONFLICT 58 58 G->D (IN REF. 2).
FT  CONFLICT 79 79 MISSING (IN REF. 2).
SQ  SEQUENCE 303 AA; 33457 MW; 809ADE0CD090023D CRC64;

Query Match      2.1%; Score 11; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  2 GHHHHHHHHH 12
DB  67 GHHHHHHHHH 77

RESULT 5
MOX2_MOUSE      STANDARD:      PRT:      303 AA.
AC  P32443;
DT  01-OCT-1993 (Rel. 27, Created)
DT  01-OCT-1993 (Rel. 27, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last annotation update)
DE  HOMEBOX PROTEIN MOX-2.
GN  MOX2 OR MOX2 OR GAX.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

RP SEQUENCE FROM N.A.
 RX MEDLINE: 9320199.
 RA Candia A.F., Hu J., Crosby J., Lalley P.A., Noden D., Nadeau J.H.,
 RT Wright C.V.E.;
 RT "Mox-1 and Mox-2 define a novel homeobox gene subfamily and are
 RT differentially expressed during early mesodermal patterning in mouse
 RT embryos.";
 RT Development 116:1123-1136(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94232829.
 RA Candia A.F., Kovalik J.-P., Wright C.V.E.;
 RT "Amino acid sequence of Mox-2 and comparison to its Xenopus and rat
 RT homologs.";
 RT Nucleic Acids Res. 21:4982-4982(1993).
 RN [3]
 RP SEQUENCE OF 1-11 FROM N.A.
 RX MEDLINE: 95349593.
 RA Andres V., Fisher S., Wearsch P., Walsh K.;
 RT "Regulation of Gax homeobox gene transcription by a combination of
 RT positive factors including myocyte-specific enhancer factor 2.";
 RT Mol. Cell. Biol. 15:4272-4281(1995).
 CC -1- FUNCTION: ROLE IN MESODERM INDUCTION AND ITS EARLIEST REGIONAL
 CC SPECIFICATION, SOMITOGENESIS, AND MYOGENIC AND SCLEROTOMAL
 CC DIFFERENTIATION. MAY HAVE A REGULATORY ROLE WHEN QUIESCENT
 CC VASCULAR SMOOTH MUSCLE CELLS REENTER THE CELL CYCLE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- DEVELOPMENTAL STAGE: IT IS NOT EXPRESSED BEFORE 8-8.5 DAYS POST
 CC COITUM. AT 8-8.5 D.P.C. IT IS FOUND ON THE ENTIRE EPITHELIUM OF
 CC THE SOMITE. AT 9.5 D.P.C. ITS EXPRESSION IS RESTRICTED TO THE
 CC SCLEROTOME. AT 10.5 D.P.C. IT IS FOUND IN SCLEROTOMALLY DERIVED
 CC CELLS INCLUDING THE VERTEBRAL AND COSTAL PRECURSORS.
 CC -1- SIMILARITY: WITH OTHER HOMEBOX PROTEINS.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: Z16406; CAA7899.1; -
 DR EMBL: S79168; CAB33758.1; -
 DR PIR: S41779; S41779.
 DR HSSP: P02833; ISAN.
 DR MGD: MGI:103219; MEOX2.
 DR PFAM: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEBOX.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS0071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.
 FT DOMAIN 42 47
 FT DOMAIN 68 79 POLY-SER.
 FT DOMAIN 80 85 POLY-HIS.
 FT DOMAIN 86 85 POLY-GLN.
 FT DOMAIN 63 85 GLN/HIS-RICH (OPA-REPEAT).
 FT DNA_BIND 186 245 HOMEBOX.
 FT SEQUENCE 303 AA; 33506 MW; 41BD05FC39AA4427 CRC64;
 SQ
 Query Match 2.1%; Score 11; DB 1; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0079;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HOMEBOX PROTEIN MOX-2 (GROWTH ARREST-SPECIFIC HOMEBOX).
 GN MEOX2 OR MOX2 OR MOX-2 OR GAX.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 93269321.
 RA Gorski D.H., Lepage D.F., Patel C.V., Copeland N.G., Jenkins N.A.,
 RT Walsh K.;
 RT "Molecular cloning of a diverged homeobox gene that is rapidly down-
 RT regulated during the G0/G1 transition in vascular smooth muscle
 RT cells.";
 RT Mol. Cell. Biol. 13:3722-3733(1993).
 RN [2]
 RP REVISIONS.
 RA Walsh K.;
 RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ROLE IN MESODERM INDUCTION AND ITS EARLIEST REGIONAL
 CC SPECIFICATION, SOMITOGENESIS, AND MYOGENIC AND SCLEROTOMAL
 CC DIFFERENTIATION. MAY HAVE A REGULATORY ROLE WHEN QUIESCENT
 CC VASCULAR SMOOTH MUSCLE CELLS REENTER THE CELL CYCLE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: AORTA AND HEART. ALSO DETECTED IN LUNG AND
 CC KIDNEY.
 CC -1- INDUCTION: RAPIDLY AND TRANSIENTLY DOWN-REGULATED DURING THE
 CC TRANSITION FROM G0 TO G1 INDUCED BY MITOGEN STIMULATION.
 CC -1- SIMILARITY: WITH OTHER HOMEBOX PROTEINS.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: Z17223; CAA78931.1; -
 DR PIR: A48130; A48130.
 DR HSSP: P02833; ISAN.
 DR PFAM: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEBOX.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS0071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.
 FT DOMAIN 42 47
 FT DOMAIN 68 79 POLY-SER.
 FT DOMAIN 80 85 POLY-HIS.
 FT DOMAIN 86 85 POLY-GLN.
 FT DOMAIN 64 85 GLN/HIS-RICH (OPA-REPEAT).
 FT DNA_BIND 186 245 HOMEBOX.
 FT SEQUENCE 303 AA; 33605 MW; 7776642AEFA3A2E8 CRC64;
 SQ
 Query Match 2.1%; Score 11; DB 1; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0079;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
 MOX2_RAT
 ID MOX2_RAT
 AC P39020;
 STANDARD;
 PRT; 303 AA.

RESULT 7
 OTX1_HUMAN
 ID OTX1_HUMAN
 AC P32242;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE HOMEOBOX PROTEIN OTX1.
 GN OTX1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 93327763.
 RA Simeone A., Acampora D., Mallamaci A., Stornaiuolo A., D'Apice M.R.,
 RA Nigro V., Boncinelli E.;
 RT "A vertebrate gene related to orthodenticle contains a homeodomain of
 RT the bicoid class and demarcates anterior neuroectoderm in the
 RT gastrulating mouse embryo."
 RL EMBL J. 12:2735-2747(1993).
 CC -1- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
 CC THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS):
 CC 5'-TCTATGCC-3'.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN.
 CC -1- DEVELOPMENTAL STAGE: EMBRYO.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 DR HSSP: P06601; 1FJL.
 DR TRANSFAC: T02079; -.
 DR MIM: 600036; -.
 DR PFAM: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEOBOX.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS50071; HOMEOBOX_2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 FT DNA_BIND 38 97 HOMEOBOX.
 FT DOMAIN 275 301 HIS-RICH.
 SO SEQUENCE 354 AA; 37327 MW; E32C1E47464BDAA CRC64;

Query Match 2.1%; Score 11; DB 1; Length 354;
 Best Local Similarity 100.0%; Pred. No. 0.0091;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 Db 291 GHHHHHHHHH 301

RESULT 8
 OTX1_MOUSE STANDARD: PRT; 355 AA.
 AC P80205;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HOMEOBOX PROTEIN OTX1.
 GN OTX1 OR OTX-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6;
 RX MEDLINE: 93327763.
 RA Simeone A., Acampora D., Mallamaci A., Stornaiuolo A., D'Apice M.R.,
 RA Nigro V., Boncinelli E.;
 RT "A vertebrate gene related to orthodenticle contains a homeodomain of
 RT the bicoid class and demarcates anterior neuroectoderm in the
 RT gastrulating mouse embryo."
 RL EMBL J. 12:2735-2747(1993).
 CC [2]
 RP SEQUENCE OF 31-98 FROM N.A.
 RX MEDLINE: 92357147.
 RA Simeone A., Acampora D., Gulisano M., Stornaiuolo A., Boncinelli E.;
 RT "Nested expression domains of four homeobox genes in developing
 RT rostral brain."
 RL Nature 358:687-690(1992).

CC -1- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
 CC THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS):
 CC 5'-TCTATGCC-3'.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: BRAIN: RESTRICTED REGIONS OF THE DEVELOPING
 CC ROSTRAL BRAIN INCLUDING THE PRESUMPTIVE CEREBRAL CORTEX AND
 CC OLFACTORY BULBS; EXPRESSED IN THE DEVELOPING OLFACTORY, AURICULAR
 CC AND OCULAR SYSTEMS, INCLUDING THE COVERING OF THE OPTIC NERVE.
 CC -1- DEVELOPMENTAL STAGE: EMBRYO.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X68883; CAA48754.1; -.
 DR PIR: S35345; S35345.
 DR HSSP: P06601; 1FJL.
 DR TRANSFAC: T02080; -.
 DR MGI: 97450; OTX1.
 DR PFAM: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEOBOX.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS50071; HOMEOBOX_2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 FT DNA_BIND 38 97 HOMEOBOX.
 FT DOMAIN 275 302 HIS-RICH.
 SO SEQUENCE 355 AA; 37531 MW; DA5136D9DAFC948F CRC64;

Query Match 2.1%; Score 11; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.0091;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 Db 291 GHHHHHHHHH 301

RESULT 9
 OTX1_RAT STANDARD: PRT; 355 AA.
 AC Q63410; Q64203;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HOMEOBOX PROTEIN OTX1.
 GN OTX1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE: 95016961.
 RA Frantz G.D., Weimann J.M., Levin M.E., McConnell S.K.;
 RT "Otx1 and Otx2 define layers and regions in developing cerebral
 RT cortex and cerebellum."
 RL J. Neurosci. 14:5725-5740(1994).
 CC [2]
 RP SEQUENCE OF 176-323 FROM N.A.
 RX MEDLINE: 96108898.
 RA Robel L., Ding M., James A.J., Lin X., Simeone A., Leckman J.F.,
 RA Vaccarino F.M.;
 RT "Fibroblast growth factor 2 increases Otx2 expression in precursor
 RT cells from mammalian telencephalon."
 RL J. Neurosci. 15:7879-7891(1995).
 CC -1- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND

CC THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS): 5'-
 CC TCTATCCC-3' (BY SIMILARITY). MAY PLAY A ROLE IN THE SPECIFICATION
 CC OR DIFFERENTIATION OF NEURONS IN THE DEEP LAYERS OF THE CEREBRAL
 CC CORTX, AND ALSO IN CEREBELLAR REGIONALIZATION DURING EARLY
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: BRAIN. RESTRICTED REGIONS OF THE DEVELOPING
 CC POSTNATAL BRAIN INCLUDING THE PRESUMPTIVE CEREBRAL CORTX AND
 CC OLFACTORY BULBS. EXPRESSED IN THE DEVELOPING OLFACTORY, AUDICULAR
 CC AND OCULAR SYSTEMS, INCLUDING THE COVERING OF THE OPTIC NERVE (BY
 CC SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT THE FOREBRAIN AND
 CC MIDBRAIN DURING DEVELOPMENT. AND IN ADDITION IS ALSO SEEN IN
 CC DISCRETE SPATIAL AND TEMPORAL DOMAINS IN THE DEVELOPING CEREBRAL
 CC CORTX AND CEREBELLUM. CONFINED TO A SUBPOPULATION OF NEURONS IN
 CC LAYERS 5 AND 6 WITHIN THE ADULT CEREBRAL CORTX AND DURING
 CC DEVELOPMENT EXPRESSION IS HIGH IN THE PROGENITORS OF THESE DEEP-
 CC LAYER CELLS. EXPRESSED IN THE DEVELOPING CEREBELLUM IN SPATIALLY
 CC RESTRICTED REGIONS OF THE EXTERNAL GRANULAR LAYER.
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: L33602; AA53557.1; -;
 CC DR EMBL: S81924; CAB3393.1; -;
 CC DR HSP: P06601; IPIJL.
 CC DR PRAM: PFO0046; homeobox; 1.
 CC DR PRINTS: PR00024; HOMEOBOX.
 CC DR PROSITE: PS00027; HOMEOBOX_1; 1.
 CC DR PROSITE: PS50071; HOMEOBOX_2; 1.
 CC KM Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 CC FT DNA_BIND 38 97
 CC FT DOMAIN 275 302 HIS-RICH.
 CC FT CONFLICT 219 219 T -> A (IN REF. 2).
 CC FT CONFLICT 229 229 R -> G (IN REF. 2).
 CC SQ SEQUENCE 355 AA; 37602 MW; C875871723D0B876 CRC64;
 CC
 CC Query Match 2.1%; Score 11; DB 1; Length 355;
 CC Best Local Similarity 100.0%; Pred. No. 0.0091;
 CC Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 GHHNNHHNNH 12
 CC DB 291 GHHNNHHNNH 301
 CC
 CC RESULT 10
 CC ID TY1_HUMAN STANDARD; PRT; 414 AA.
 CC AC P25490; Q14935;
 CC DT 01-MAY-1992 (Rel. 22, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
 CC DE TRANSCRIPTIONAL REPRESSOR PROTEIN TY1 (YIN AND YANG 1) (YY-1) (DELTA
 CC DE TRANSCRIPTION FACTOR) (NF-E1).
 CC GN TY1.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 CC RN [1]
 CC RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 CC RX MEDLINE; 92005716.
 CC RT Shi Y., Seto E., Chang L.-S., Shenk T.;
 CC RT "Transcriptional repression by YY1, a human G1-Kruppel-related
 CC RT protein, and relief of repression by adenovirus E1a protein.";

RL Cell 67:377-386(1991).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE-FORESKIN;
 RC MEDLINE; 92052179.
 RX Park K., Atchison M.;
 RA "Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta),
 RT that binds to the immunoglobulin kappa 3' enhancer and the
 RT immunoglobulin heavy-chain mu E1 site.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9804-9808(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Whitson R.H., Huang T., Dang J., Itakura K.;
 RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SUBCELLULAR LOCATION.
 RX MEDLINE; 98152931.
 RA McNeil S., Guo B., Stein J.L., Iian J.B., Bushmeyer S., Seto E.,
 RA Atchison M.L., Pennan S., van Wijnen A.J., Stein G.S.;
 RT "Targeting of the YY1 transcription factor to the nucleolus and the
 RT nuclear matrix in situ: the C-terminus is a principal determinant for
 RT nuclear trafficking.";
 RL J. Cell. Biochem. 68:500-510(1998).
 RN [5]
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 304-414.
 RX MEDLINE; 97098436.
 RA Houbavly H.B., Ushveva A., Shenk T., Burley S.K.;
 RT "Cocrystal structure of YY1 bound to the adeno-associated virus P5
 RT initiator.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13577-13582(1996).
 RN [6]
 RN STRUCTURE BY NMR OF 353-379.
 RX MEDLINE; 98308000.
 RA Viles J.H., Patel S.U., Mitchell J.B.O., Moody C.M., Justice D.E.,
 RA Updegrink J., Doyle P.M., Harris C.J., Sadler P.J., Thornton J.M.;
 RT "Design, synthesis and structure of a zinc finger with an artificial
 RT beta-turn.";
 RL J. Mol. Biol. 279:973-986(1998).
 CC -1- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
 CC POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
 CC VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION
 CC START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
 CC DIFFERENTIATION. THE FUNCTION OF YY1 AS AN ACTIVATOR OR A
 CC REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. FOR
 CC EXAMPLE IT ACTS AS AN ACTIVATOR IN ITS PRESENCE.
 CC PROTEIN BUT AS AN ACTIVATOR IN ITS PRESENCE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
 CC -1- SIMILARITY: BELONGS TO THE KRUPPEL SUBFAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: M77698; AA59467.1; -;
 CC DR EMBL: M76541; AA58926.1; -;
 CC DR EMBL: Z14077; CAH78435.1; -;
 CC DR PIR: A40350; A40350.
 CC DR PDB: 1UBD; 23-DEC-96.
 CC DR PDB: 1ZNM; 01-APR-98.
 CC DR TRANSPAC: T00915; -;
 CC DR MTM: 600013; -;
 CC DR PRAM: PFO0096; Zf-C2H2; 4.
 CC DR PRINTS: PR00048; ZINC-FINGER.
 CC DR PROSITE: PS00028; ZINC_FINGER_C2H2; 4.
 CC KM Transcription regulation; Repressor; Activator; Nuclear protein;
 CC RT Zinc-finger; Metal-binding; DNA-binding; Repeat; 3D-structure.
 FT DOMAIN 43 53
 FT DOMAIN 54 69 G1Y-RICH.

```

FT DOMAIN 70 80 POLY-HIS.
FT DOMAIN 159 170 GLY/SER-RICH.
FT FT 296 407 4 C2H2-TYPE ZINC-FINGERS.
FT ZN_FING 296 330 C2H2-TYPE.
FT ZN_FING 325 347 C2H2-TYPE.
FT ZN_FING 353 377 C2H2-TYPE.
FT ZN_FING 383 407 C2H2-TYPE.
FT ZN_FING 383 407 C2H2-TYPE.
FT DOMAIN 333 371 INVOLVED IN NUCLEAR MATRIX ASSOCIATION.
FT DOMAIN 333 371 INVOLVED IN REPRESSION OF ACTIVATED
TRANSCRIPTION.
FT DOMAIN 371 397 INVOLVED IN MASKING TRANSCRIPTION.
FT CONFLICT 65 65 H -> R (IN REF. 2).
FT CONFLICT 196 196 G -> R (IN REF. 1).
SQ SEQUENCE 414 AA: 44712 MW: 058C05A0AD2D04E6 CRC64;

Query Match 2.1%; Score 11; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
Db 69 GHHHHHHHHH 79

RESULT 11
TY1_MOUSE STANDARD; PRT; 414 AA.
ID TY1_MOUSE
AC 000899;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TRANSCRIPTIONAL REPRESSOR PROTEIN YY1 (YY1 AND YANG 1) (YY-1) (DELTA
DE TRANSCRIPTION FACTOR) (NF-E1) (UCR-MOTIF DNA-BINDING PROTEIN).
GN YY1 OR UCRBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 92107191.
RA Flanagan J.R., Becker K.G., Enlist D.L., Gleason S.L., Driggers P.H.,
RA Levi B.-Z., Appella E., Ozato R.;
RT "Cloning of a negative transcription factor that binds to the
RT upstream conserved region of Moloney murine leukemia virus.";
RT Mol. Cell. Biol. 12:38-44(1992).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE: 93296177.
RA Safran G., Perry R.P.;
RT "Characterization of the mouse gene that encodes the delta/YY1/NF-
RT EI/OCBP transcription factor.";
RT Proc. Natl. Acad. Sci. U.S.A. 90:5559-5563(1993).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE: 92052178.
RA Haribaran N., Kelley D.E., Perry R.P.;
RT "Delta, a transcription factor that binds to downstream elements in
RT several polymerase II promoters, is a functionally versatile zinc
RT finger protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9799-9803(1991).
CC -1- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
CC POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
CC VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION
CC START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
CC DIFFERENTIATION. THE FUNCTION OF YY1 AS AN ACTIVATOR OR A
CC REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. BINDS TO
CC THE UPSTREAM CONSERVED REGION (UCR) (5'-GCCCATTTT-3') OF MOLONEY
CC MURINE LEUKEMIA VIRUS (MDLV).
CC SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-

```

```

CC CC FINGER PROTEINS.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC CC or send an email to license@sdb.ch).
CC CC -----
DR EMBL: M73963; AAA40522.1; -
DR EMBL: L13968; AAA40477.1; -
DR EMBL: L13969; AAA40477.1; JOINED.
DR EMBL: L13965; AAA40477.1; JOINED.
DR EMBL: L13966; AAA40477.1; JOINED.
DR EMBL: L13967; AAA40477.1; JOINED.
DR EMBL: M74580; AAA37521.1; -
DR HSSP: P23480; 1UBD.
DR MGD: MGI:99150; YY1.
DR PFAM: PF000096; zf-C2H2; 4.
DR PRINTS: PR00048; ZINC_FINGER.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 4.
KM Transcription regulation. Repressor; Activator; Nuclear protein;
KM Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT DOMAIN 43 53 GLY/SER-RICH.
FT DOMAIN 54 70 POLY-HIS.
FT DOMAIN 71 82 GLY/SER-RICH.
FT DOMAIN 161 170 POLY-HIS.
FT DOMAIN 286 407 4 C2H2-TYPE ZINC-FINGERS.
FT ZN_FING 296 330 C2H2-TYPE.
FT ZN_FING 325 347 C2H2-TYPE.
FT ZN_FING 353 377 C2H2-TYPE.
FT ZN_FING 383 407 C2H2-TYPE.
FT ZN_FING 383 407 C2H2-TYPE.
FT DOMAIN 257 341 INVOLVED IN NUCLEAR MATRIX ASSOCIATION
(BY SIMILARITY).
FT DOMAIN 333 371 INVOLVED IN REPRESSION OF ACTIVATED
TRANSCRIPTION (BY SIMILARITY).
FT DOMAIN 371 397 INVOLVED IN MASKING TRANSCRIPTION.
FT CONFLICT 219 219 F -> S (IN REF. 3).
FT CONFLICT 375 375 R -> G (IN REF. 3).
SQ SEQUENCE 414 AA: 44717 MW: C01237828B8E984F9 CRC64;

Query Match 2.1%; Score 11; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
Db 70 GHHHHHHHHH 80

RESULT 12
ZIC3_MOUSE STANDARD; PRT; 466 AA.
ID ZIC3_MOUSE
AC 062521;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ZINC FINGER PROTEIN ZIC3 (ZINC FINGER PROTEIN OF THE CEREBELLUM 3).
GN ZIC3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-CEREBELLUM;
RX MEDLINE: 96132843.
RA Aruga J., Nagai T., Tokuyama T., Hayashizaki Y., Okazaki Y.,
RA Chapman V.M., Mikoshiba K.;
RT "The mouse zic gene family. Homologues of the Drosophila pair-rule
RT "gene odd-paired".";

```


DR EMBL: L78130; AAB48642.1; --
 DR HSSP: P24941; 1A01.
 DR PFAM: PF00069; PKinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 KW transferase; Serine/threonine-protein kinase; ATP-binding;
 KM phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 6 17
 FT NP_BIND 39 290
 FT BINDING 45 53
 FT ACT_SITE 68 68
 FT MOD_RES 161 161
 FT ACT_SITE 194 194
 SQ SEQUENCE 611 AA; 70049 MW; 89E17812A490CCD0 CRC64; BY SIMILARITY).

Query Match 2.1%; Score 11; DB 1; Length 611;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HHHHHHHHS 13
 |||||
 Db 8 HHHHHHHHS 18

RESULT 15
 SNF1_CANTR
 ID SNF1_CANTR STANDARD; PRT; 619 AA.
 AC 094168;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CARBON CATABOLITE DEREPRESSING PROTEIN KINASE (EC 2.7.1.-).
 GN SNF1
 OS Candida tropicalis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC anamorphic Saccharomycetales; Candida.
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Kanai T., Ogawa K., Ueda M., Tanaka A.;
 RT "Genetic evaluation of the function of SNF1 in Candida tropicalis";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
 INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
 PROTEIN SNF4. COULD PHOSPHORYLATES CAP8 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
 SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB024535; BAA75889.1; --
 DR PFAM: PF00069; PKinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 KW transferase; Serine/threonine-protein kinase; ATP-binding;
 KM phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 6 17
 FT NP_BIND 39 290
 FT BINDING 45 53
 FT ACT_SITE 68 68
 FT MOD_RES 161 161
 FT ACT_SITE 194 194
 SQ SEQUENCE 619 AA; 70323 MW; 0FCF1C3DCE706D7 CRC64; BY SIMILARITY).

Query Match 2.1%; Score 11; DB 1; Length 619;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HHHHHHHHS 13
 |||||
 Db 20 HHHHHHHHS 30

Search completed: August 22, 2000, 17:47:11
 Job time: 310 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2000, 17:41:11 ; Search time 46.55 seconds
(without alignments)
776.006 Million cell updates/sec

Title: US-08-699-716a-2
Perfect score: 521
Sequence: 1 MGNHHHHHHHSSGHIDD.....RFIOKYDSVMORLLDITSGK 521

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 225878 seqs, 69334122 residues

Word size : 0

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

SPREMBL_12:*

1: sp.archaea:*

2: sp.bacteria:*

3: sp.fungi:*

4: sp.human:*

5: sp.invertebrate:*

6: sp.mammal:*

7: sp.mhc:*

8: sp.organelle:*

9: sp.phage:*

10: sp.plant:*

11: sp.podent:*

12: sp.virus:*

13: sp.vertibrate:*

14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	23	46.6	326	2	068697	068697 yersinia pe
2	104	20.0	334	2	056853	056853 yersinia en
3	104	20.0	334	2	056896	056896 yersinia en
4	96	18.4	324	2	056846	056846 yersinia en
5	96	18.4	324	2	056879	056879 yersinia en
6	93	17.9	325	2	056880	056880 yersinia en
7	68	13.1	324	2	087495	087495 yersinia en
8	67	12.9	324	2	056882	056882 yersinia en
9	33	3.3	168	4	016256	016256 homo sapien
10	17	2.5	294	2	030527	030527 pseudomonas
11	12	2.3	516	13	042186	042186 brachydanio
12	11	2.3	230	4	075263	075263 homo sapien
13	11	2.1	420	3	042980	042980 schizosacch
14	11	2.1	427	3	P78807	P78807 schizosacch
15	11	2.1	467	13	09YGC6	09YGC6 xenopus lae
16	11	2.1	911	5	097423	097423 drosophila
17	11	2.1	1180	5	024163	024163 drosophila
18	10	1.9	53	14	055244	055244 plasmod pet
19	10	1.9	202	13	Q90691	Q90691 gallus galli

20	10	1.9	204	6	Q28555	Q28555 ovis aries
21	10	1.9	309	5	Q18751	Q18751 caenorhabd
22	10	1.9	321	10	Q23891	Q23891 oryza sativ
23	10	1.9	391	10	Q23038	Q23038 arabidopsis
24	10	1.9	407	10	Q920A3	Q920A3 arabidopsis
25	10	1.9	443	4	Q14979	Q14979 homo sapien
26	10	1.9	465	11	P70512	P70512 rattus norv
27	10	1.9	465	11	008755	008755 mus musculu
28	10	1.9	491	11	008755	008755 rattus norv
29	10	1.9	550	11	070369	070369 rattus norv
30	10	1.9	551	11	055187	055187 mus musculu
31	10	1.9	558	4	000257	000257 homo sapien
32	10	1.9	588	12	066043	066043 simian herp
33	10	1.9	604	13	Q42472	Q42472 xenopus lae
34	10	1.9	716	5	097216	097216 leishmania
35	10	1.9	791	11	070368	070368 rattus norv
36	10	1.9	884	5	061543	061543 drosophila
37	10	1.9	905	5	Q26344	Q26344 aplysia sp.
38	10	1.9	977	11	P97838	P97838 rattus norv
39	10	1.9	980	4	095387	095387 homo sapien
40	10	1.9	1100	4	094875	094875 homo sapien
41	10	1.9	1172	10	064474	064474 arabidopsis
42	10	1.9	1196	11	Q35413	Q35413 rattus norv
43	10	1.9	1212	5	Q24523	Q24523 drosophila
44	10	1.9	1273	4	060316	060316 homo sapien
45	10	1.9	1291	5	077261	077261 drosophila
46	10	1.9	1622	5	065550	065550 cryptospori
47	9	1.7	44	12	Q65825	Q65825 undulatifile
48	9	1.7	136	11	008657	008657 rattus norv
49	9	1.7	216	13	Q9W7C7	Q9W7C7 oryzias lat
50	9	1.7	229	10	065124	065124 dicentra ex
51	9	1.7	311	13	Q90370	Q90370 coturnix co
52	9	1.7	311	13	Q90888	Q90888 gallus galli
53	9	1.7	324	13	Q9W6B1	Q9W6B1 brachydanio
54	9	1.7	348	10	Q94457	Q94457 arabidopsis
55	9	1.7	356	13	Q73679	Q73679 brachydanio
56	9	1.7	394	5	Q9YXQ3	Q9YXQ3 anopheles g
57	9	1.7	497	13	Q9Y1B7	Q9Y1B7 xenopus lae
58	9	1.7	501	13	Q93487	Q93487 xenopus lae
59	9	1.7	501	13	Q91689	Q91689 xenopus lae
60	9	1.7	509	5	Q94888	Q94888 drosophila
61	9	1.7	509	5	Q9Y0H8	Q9Y0H8 drosophila
62	9	1.7	515	11	054949	054949 mus musculu
63	9	1.7	522	5	Q9XTM5	Q9XTM5 drosophila
64	9	1.7	522	5	Q9XTK9	Q9XTK9 drosophila
65	9	1.7	533	4	Q95409	Q95409 homo sapien
66	9	1.7	627	4	Q94825	Q94825 homo sapien
67	9	1.7	678	5	Q94736	Q94736 stomoxys ca
68	9	1.7	693	5	096680	096680 schizosacch
69	9	1.7	794	3	Q42890	Q42890 schizosacch
70	9	1.7	808	4	Q15083	Q15083 homo sapien
71	9	1.7	1116	3	Q99247	Q99247 saccharomyc
72	9	1.7	1155	4	Q94770	Q94770 homo sapien
73	9	1.7	1560	4	060275	060275 homo sapien
74	9	1.7	2243	4	Q9Y5T3	Q9Y5T3 homo sapien
75	9	1.7	2250	4	Q9Y5T2	Q9Y5T2 homo sapien
76	9	1.7	2288	11	Q9WUT8	Q9WUT8 rattus norv
77	9	1.7	2295	11	Q9WUT2	Q9WUT2 mus musculu
78	9	1.7	2353	4	Q95802	Q95802 homo sapien
79	9	1.5	83	5	Q20689	Q20689 caenorhabd
80	8	1.5	102	5	Q94189	Q94189 caenorhabd
81	8	1.5	115	1	09YAJ2	09YAJ2 aeropyrum p
82	8	1.5	140	5	Q26056	Q26056 plasmodium
83	8	1.5	183	5	Q20741	Q20741 caenorhabd
84	8	1.5	208	2	084365	084365 chlamydia t
85	8	1.5	216	11	061099	061099 mus musculu
86	8	1.5	216	11	064279	064279 mus musculu
87	8	1.5	216	11	P97832	P97832 rattus norv
88	8	1.5	250	5	P90780	P90780 caenorhabd
89	8	1.5	274	10	Q9XGS2	Q9XGS2 gylcine max
90	8	1.5	286	13	057342	057342 coturnix co
91	8	1.5	286	13	Q42290	Q42290 gallus galli
92	8	1.5	292	13	Q91294	Q91294 rana catesb

93	1.5	323	4	09Y503	09Y593 homo sapien	166	7	1.3	290	5	045469	045469 caenorhabdi
94	1.5	326	13	093336	093336 xenopus lae	167	7	1.3	294	2	P94395	P94395 bacillus su
95	1.5	358	12	010418	010418 human cytom	168	7	1.3	295	10	022807	022807 arabidopsis
96	1.5	358	12	069215	069215 human cytom	169	7	1.3	295	10	09ZV85	09ZV85 arabidopsis
97	1.5	370	5	024150	024150 drosophila	170	7	1.3	305	5	P91314	P91314 caenorhabdi
98	1.5	401	3	006178	006178 saccharomye	171	7	1.3	306	11	035449	035449 mus musculu
99	1.5	409	3	012128	012128 saccharomye	172	7	1.3	314	2	025229	025229 helicobacte
100	1.5	423	12	010417	010417 human cytom	173	7	1.3	314	5	09Z1X8	09Z1X8 helicobacte
101	1.5	484	11	092206	092206 mus musculu	174	7	1.3	314	5	021760	021760 caenorhabdi
102	1.5	485	4	095948	095948 homo sapien	175	7	1.3	314	10	09ZUY7	09ZUY7 arabidopsi
103	1.5	559	5	09XUL9	09XUL9 caenorhabdi	176	7	1.3	322	10	064452	064452 nicotiana p
104	1.5	576	5	024180	024180 drosophila	177	7	1.3	329	2	09XC70	09XC70 streptomye
105	1.5	684	12	P87888	P87888 human cytom	178	7	1.3	342	10	065004	065004 lycopersico
106	1.5	704	5	09XTN4	09XTN4 drosophila	179	7	1.3	342	10	042658	042658 caenorhabdi
107	1.5	895	4	09Y2X9	09Y2X9 homo sapien	180	7	1.3	343	5	001575	001575 caenorhabdi
108	1.5	1133	5	09S050	09S050 tetrahymena	181	7	1.3	348	8	099921	099921 cyprinella
109	1.5	1172	4	09Y4F2	09Y4F2 homo sapien	182	7	1.3	348	8	099922	099922 cyprinella
110	1.5	1203	5	022484	022484 caenorhabdi	183	7	1.3	373	10	081369	081369 arabidopsis
111	1.5	1441	5	096957	096957 drosophila	184	7	1.3	377	11	008949	008949 rattus norv
112	1.5	1557	5	096652	096652 drosophila	185	7	1.3	384	10	082136	082136 pisum sativ
113	1.5	1597	5	061346	061346 drosophila	186	7	1.3	394	2	P96877	P96877 thermotoga
114	1.5	1929	5	093637	093637 caenorhabdi	187	7	1.3	398	2	09WZ57	09WZ57 mycobacteri
115	1.5	2416	5	P91667	P91667 drosophila	188	7	1.3	406	2	047280	047280 escherichia
116	1.5	5327	5	076891	076891 drosophila	189	7	1.3	411	2	P78701	P78701 metarhizium
117	1.3	47	12	09WH80	09WH80 chimpanzee	190	7	1.3	411	2	09Z361	09Z361 streptomye
118	1.3	77	5	020630	020630 caenorhabdi	191	7	1.3	425	5	076831	076831 caenorhabdi
119	1.3	83	2	P76094	P76094 escherichia	192	7	1.3	434	10	023513	023513 arabidopsis
120	1.3	86	13	042437	042437 notophthalm	193	7	1.3	444	3	091765	091765 xenopus lae
121	1.3	127	2	092385	092385 synecococc	194	7	1.3	446	6	097727	097727 sus scrofa
122	1.3	136	5	093827	093827 caenorhabdi	195	7	1.3	451	10	082690	082690 lepicloclis
123	1.3	142	2	09X781	09X781 streptomye	196	7	1.3	456	2	054101	054101 saccharopo
124	1.3	146	2	092FC2	092FC2 mycobacteri	197	7	1.3	456	2	048873	048873 lactobacill
125	1.3	148	2	034913	034913 enterococcu	198	7	1.3	458	2	032907	032907 mycobacteri
126	1.3	149	2	P97026	P97026 bacillus 11	199	7	1.3	466	2	09ZDM0	09ZDM0 rickettsia
127	1.3	154	4	09Y3C5	09Y3C5 homo sapien	200	7	1.3	470	2	004317	004317 arabidopsis
128	1.3	158	5	094577	094577 helioicidari	201	7	1.3	483	2	054179	054179 streptomye
129	1.3	158	5	046248	046248 drosophila	202	7	1.3	499	5	044658	044658 caenorhabdi
130	1.3	159	5	046258	046258 drosophila	203	7	1.3	504	12	081961	081961 human papil
131	1.3	159	12	09YMT6	09YMT6 lymantria d	204	7	1.3	513	2	032561	032561 escherichia
132	1.3	182	1	059259	059259 pyrococcus	205	7	1.3	513	10	040029	040029 hordium vul
133	1.3	185	1	000837	000837 leishmania	206	7	1.3	513	10	040030	040030 hordium vul
134	1.3	187	10	049684	049684 arabidopsis	207	7	1.3	513	10	081992	081992 hordium vul
135	1.3	192	5	046232	046232 drosophila	208	7	1.3	513	10	080567	080567 arabidopsis
136	1.3	192	5	046260	046260 drosophila	209	7	1.3	532	2	069975	069975 streptomye
137	1.3	193	5	046242	046242 drosophila	210	7	1.3	533	6	097623	097623 canis famli
138	1.3	195	5	046262	046262 drosophila	211	7	1.3	533	13	09YGX2	09YGX2 gallus gall
139	1.3	196	5	046234	046234 drosophila	212	7	1.3	539	12	09YF0S0	09YF0S0 human calic
140	1.3	196	5	046256	046256 drosophila	213	7	1.3	539	12	09WRZ1	09WRZ1 normal-walk-lik
141	1.3	198	5	046238	046238 drosophila	214	7	1.3	539	12	09WRZ0	09WRZ0 normal-walk-lik
142	1.3	198	5	046240	046240 drosophila	215	7	1.3	539	12	09WRZ0	09WRZ0 normal-walk-lik
143	1.3	198	5	046244	046244 drosophila	216	7	1.3	539	12	09WRZ0	09WRZ0 normal-walk-lik
144	1.3	199	5	046246	046246 drosophila	217	7	1.3	539	12	09WRZ0	09WRZ0 normal-walk-lik
145	1.3	201	2	050412	050412 magnetospir	218	7	1.3	539	12	09WRZ0	09WRZ0 normal-walk-lik
146	1.3	209	12	084640	084640 parametium	219	7	1.3	539	12	09WRZ0	09WRZ0 normal-walk-lik
147	1.3	211	4	075333	075333 homo sapien	220	7	1.3	540	12	09WRZ0	09WRZ0 normal-walk-lik
148	1.3	213	4	096004	096004 homo sapien	221	7	1.3	540	12	09WRZ0	09WRZ0 normal-walk-lik
149	1.3	219	12	055579	055579 leucania se	222	7	1.3	550	5	021540	021540 caenorhabdi
150	1.3	235	5	076647	076647 caenorhabdi	223	7	1.3	550	5	045494	045494 caenorhabdi
151	1.3	235	5	054779	054779 streptococc	224	7	1.3	587	5	021432	021432 caenorhabdi
152	1.3	244	10	064461	064461 brassica na	225	7	1.3	587	4	060324	060324 homo sapien
153	1.3	253	3	093949	093949 candida alb	226	7	1.3	591	2	092493	092493 flavobacter
154	1.3	253	3	094052	094052 caenorhabdi	227	7	1.3	601	5	077459	077459 drosophila
155	1.3	256	4	013722	013722 homo sapien	228	7	1.3	617	12	083641	083641 measles vir
156	1.3	256	4	015747	015747 dictyosteli	229	7	1.3	619	10	004316	004316 arabidopsis
157	1.3	259	2	066901	066901 aquifex aeo	230	7	1.3	620	3	012100	012100 saccharomye
158	1.3	259	10	09ZTC8	09ZTC8 arabidopsis	231	7	1.3	620	10	023223	023223 arabidopsis
159	1.3	263	5	020288	020288 caenorhabdi	232	7	1.3	622	13	098985	098985 rana rugosa
160	1.3	263	12	065175	065175 african swi	233	7	1.3	643	6	097726	097726 sus scrofa
161	1.3	273	5	076688	076688 caenorhabdi	234	7	1.3	647	3	014273	014273 schizosacch
162	1.3	276	13	006962	006962 salmonella	235	7	1.3	670	2	P74816	P74816 sphingomon
163	1.3	276	13	P79788	P79788 gallus gall	236	7	1.3	685	4	060679	060679 homo sapien
164	1.3	281	2	067136	067136 aquifex aeo	237	7	1.3	688	5	022402	022402 caenorhabdi
165	1.3	284	2	054187	054187 streptomye	238	7	1.3	701	5	019460	019460 caenorhabdi

239	7	1.3	703	11	061230	061230 mus musculu	312	6	1.2	79	2	Q9x214	Q9x214 thermotoga
240	7	1.3	732	4	014586	014586 homo sapien	313	6	1.2	79	10	Q42170	Q42170 arabidopsis
241	7	1.3	733	4	028972	028972 archaeoglob	314	6	1.2	81	2	054558	054558 streptococ
242	7	1.3	793	4	092119	092119 homo sapien	315	6	1.2	81	12	090930	090930 human immun
243	7	1.3	794	2	09x212	09x212 thermotoga	316	6	1.2	85	6	P79306	P79306 sus scrofa
244	7	1.3	795	5	09y115	09y115 dictyostell	317	6	1.2	85	12	Q9MBE1	Q9MBE1 human immun
245	7	1.3	798	11	088411	088411 mus musculu	318	6	1.2	86	2	Q48627	Q48627 leuconostoc
246	7	1.3	799	4	094874	094874 homo sapien	319	6	1.2	86	12	Q9YWS4	Q9YWS4 human immun
247	7	1.3	807	12	Q9YTO5	Q9YTO5 ateline herp	320	6	1.2	86	12	Q9YWS3	Q9YWS3 human immun
248	7	1.3	808	12	069320	069320 gallid herp	321	6	1.2	87	5	001463	001463 caenorhabd1
249	7	1.3	820	4	000236	000236 homo sapien	322	6	1.2	87	12	055317	055317 human immun
250	7	1.3	829	5	024183	024183 drosophila	323	6	1.2	88	2	066152	066152 actinobacil
251	7	1.3	843	4	09Y2T1	09Y2T1 homo sapien	324	6	1.2	88	4	Q94795	Q94795 homo sapien
252	7	1.3	844	12	041789	041789 human immun	325	6	1.2	88	12	Q9YWS5	Q9YWS5 human immun
253	7	1.3	886	13	042588	042588 xenopus lae	326	6	1.2	88	12	Q9WCW4	Q9WCW4 human immun
254	7	1.3	892	2	055676	055676 synchocyst	327	6	1.2	89	2	P74811	P74811 synchocyst
255	7	1.3	904	11	008971	008971 mus musculu	328	6	1.2	89	2	Q9ZH04	Q9ZH04 yersinia pe
256	7	1.3	932	5	P89499	P89499 saccharomyc	329	6	1.2	89	2	Q9X0Q8	Q9X0Q8 thermotoga
257	7	1.3	1001	5	022534	022534 caenorhabd1	330	6	1.2	89	12	Q78979	Q78979 human immun
258	7	1.3	1002	2	087109	087109 serralia ma	331	6	1.2	89	12	069683	069683 human immun
259	7	1.3	1041	10	023536	023536 arabidopsis	332	6	1.2	90	10	022295	022295 actinidid c
260	7	1.3	1133	5	024034	024034 drosophila	333	6	1.2	90	12	069703	069703 human immun
261	7	1.3	1133	5	077426	077426 drosophila	334	6	1.2	90	12	069652	069652 human immun
262	7	1.3	1150	5	094548	094548 drosophila	335	6	1.2	90	12	069674	069674 human immun
263	7	1.3	1176	12	Q98587	Q98587 paramecium	336	6	1.2	90	12	Q9YWS2	Q9YWS2 human immun
264	7	1.3	1189	4	043604	043604 homo sapien	337	6	1.2	90	12	Q9YWS2	Q9YWS2 human immun
265	7	1.3	1199	4	060668	060668 homo sapien	338	6	1.2	91	12	069515	069515 human herpe
266	7	1.3	1199	4	043487	043487 homo sapien	339	6	1.2	91	12	069697	069697 human immun
267	7	1.3	1289	3	008748	008748 saccharomyc	340	6	1.2	91	12	Q76757	Q76757 human immun
268	7	1.3	1314	4	Q12769	Q12769 homo sapien	341	6	1.2	91	12	Q76758	Q76758 human immun
269	7	1.3	1318	12	P90493	P90493 herpes simp	342	6	1.2	91	12	Q76759	Q76759 human immun
270	7	1.3	1361	2	086617	086617 streptomyce	343	6	1.2	91	12	Q76760	Q76760 human immun
271	7	1.3	1378	12	Q9WRS7	Q9WRS7 macaca mula	344	6	1.2	91	12	Q76761	Q76761 human immun
272	7	1.3	1402	11	0920W3	0920W3 mus musculu	345	6	1.2	91	12	Q76762	Q76762 human immun
273	7	1.3	1405	5	017909	017909 caenorhabd1	346	6	1.2	91	12	Q9WCW5	Q9WCW5 human immun
274	7	1.3	1530	4	043241	043241 homo sapien	347	6	1.2	91	12	Q9XCH6	Q9XCH6 rickettsia
275	7	1.3	1531	2	084818	084818 chlamydia t	348	6	1.2	92	2	018410	018410 caenorhabd1
276	7	1.3	1705	2	005179	005179 bordetella	349	6	1.2	92	5	019331	019331 caenorhabd1
277	7	1.3	1894	11	064487	064487 mus musculu	350	6	1.2	92	5	019331	019331 caenorhabd1
278	7	1.3	1894	5	096133	096133 plasmodium	351	6	1.2	92	12	Q79228	Q79228 human immun
279	7	1.3	2117	2	Q92429	Q92429 streptomyce	352	6	1.2	92	12	Q70063	Q70063 human immun
280	7	1.3	3056	12	065852	065852 bean yellow	353	6	1.2	92	12	Q76931	Q76931 human immun
281	7	1.3	3056	12	065892	065892 bean yellow	354	6	1.2	92	12	Q76932	Q76932 human immun
282	7	1.3	3072	12	093645	093645 clover yell	355	6	1.2	92	12	Q76933	Q76933 human immun
283	7	1.3	3104	2	004846	004846 corynebacte	356	6	1.2	92	12	Q76935	Q76935 human immun
284	7	1.3	3345	5	046074	046074 drosophila	357	6	1.2	92	12	Q76938	Q76938 human immun
285	7	1.3	3413	2	054593	054593 amycolatops	358	6	1.2	92	12	Q76949	Q76949 human immun
286	7	1.3	3729	2	033956	033956 streptomyce	359	6	1.2	92	12	Q76950	Q76950 human immun
287	7	1.2	16	2	048439	048439 klebsiella	360	6	1.2	92	12	Q76951	Q76951 human immun
288	6	1.2	34	5	094583	094583 heliothis v	361	6	1.2	92	12	Q76954	Q76954 human immun
289	6	1.2	41	2	006527	006527 lactobacill	362	6	1.2	92	12	Q76955	Q76955 human immun
290	6	1.2	41	2	065824	065824 human immun	363	6	1.2	92	12	Q76956	Q76956 human immun
291	6	1.2	54	12	065826	065826 human immun	364	6	1.2	92	12	Q76958	Q76958 human immun
292	6	1.2	57	12	065114	065114 human herpe	365	6	1.2	92	12	Q76959	Q76959 human immun
293	6	1.2	57	12	069122	069122 human herpe	366	6	1.2	92	12	Q76934	Q76934 human immun
294	6	1.2	61	12	090268	090268 hepatitis b	367	6	1.2	92	12	Q76936	Q76936 human immun
295	6	1.2	62	2	034390	034390 bacillus su	368	6	1.2	92	12	Q76937	Q76937 human immun
296	6	1.2	64	10	Q40049	Q40049 hordeum vul	369	6	1.2	92	12	Q76948	Q76948 human immun
297	6	1.2	65	1	Q51794	Q51794 pyrococcus	370	6	1.2	92	12	Q76953	Q76953 human immun
298	6	1.2	66	10	Q22226	Q22226 arabidopsis	371	6	1.2	92	12	Q76953	Q76953 human immun
299	6	1.2	70	2	P71998	P71998 mycobacteri	372	6	1.2	92	12	Q76960	Q76960 human immun
300	6	1.2	70	11	088380	088380 mus musculu	373	6	1.2	92	12	Q76961	Q76961 human immun
301	6	1.2	72	2	025973	025973 helicobacte	374	6	1.2	92	12	Q9YWS1	Q9YWS1 human immun
302	6	1.2	73	9	092X34	092X34 mycobacteri	375	6	1.2	92	12	Q9WCW2	Q9WCW2 human immun
303	6	1.2	76	5	Q23550	Q23550 naegleria f	376	6	1.2	93	2	085974	085974 splingomona
304	6	1.2	76	10	065488	065488 arabidopsis	377	6	1.2	93	12	Q75328	Q75328 human immun
305	6	1.2	77	2	Q62Y18	Q62Y18 helicobacte	378	6	1.2	93	12	Q97199	Q97199 human immun
306	6	1.2	77	12	Q9WC11	Q9WC11 human immun	379	6	1.2	93	12	Q97200	Q97200 human immun
307	6	1.2	77	13	Q13141	Q13141 cyprinus ca	380	6	1.2	93	12	Q97201	Q97201 human immun
308	6	1.2	78	2	Q32805	Q32805 lactococcus	381	6	1.2	93	12	Q97202	Q97202 human immun
309	6	1.2	78	2	Q9XV55	Q9XV55 caenorhabd1	382	6	1.2	93	12	Q80338	Q80338 human immun
310	6	1.2	78	12	Q70489	Q70489 human immun	383	6	1.2	93	12	Q80331	Q80331 human immun
311	6	1.2	79	2	085256	085256 streptococ	384	6	1.2	93	12	Q80333	Q80333 human immun

385	6	1.2	93	12	Q9YWS3	Q9YWS3 human immun	458	6	1.2	112	1	027980	027980 archaeoglob
386	6	1.2	93	12	Q9WCP2	Q9WCP2 human immun	459	6	1.2	112	12	P88144	P88144 human immun
387	6	1.2	93	12	Q9WCL7	Q9WCL7 human immun	460	6	1.2	113	13	042461	042461 squallus aca
388	6	1.2	93	12	Q9WCK8	Q9WCK8 human immun	461	6	1.2	113	13	042462	042462 squallus aca
389	6	1.2	93	12	Q9WCH6	Q9WCH6 human immun	462	6	1.2	113	13	012967	012967 fuigu rubrip
390	6	1.2	94	12	Q97198	Q97198 human immun	463	6	1.2	114	5	046154	046154 plutella xy
391	6	1.2	94	12	Q97204	Q97204 human immun	464	6	1.2	114	5	046155	046155 plutella xy
392	6	1.2	94	12	Q9WCT4	Q9WCT4 human immun	465	6	1.2	114	5	096793	096793 myzus persi
393	6	1.2	94	12	Q9WCS6	Q9WCS6 human immun	466	6	1.2	115	5	017259	017259 brachionus
394	6	1.2	94	12	Q9WCS3	Q9WCS3 human immun	467	6	1.2	115	12	Q9YL39	Q9YL39 human immun
395	6	1.2	94	12	Q9WCR4	Q9WCR4 human immun	468	6	1.2	115	12	Q9YKGS	Q9YKGS human immun
396	6	1.2	94	12	Q9WCM7	Q9WCM7 human immun	469	6	1.2	116	2	P76837	P76837 escherichia
397	6	1.2	94	12	Q9WCH0	Q9WCH0 human immun	470	6	1.2	116	12	067741	067741 human astro
398	6	1.2	94	12	Q9WCG9	Q9WCG9 human immun	471	6	1.2	116	12	067742	067742 human astro
399	6	1.2	95	5	Q24866	Q24866 entamoeba i	472	6	1.2	117	8	020157	020157 chlorella v
400	6	1.2	95	9	Q38398	Q38398 bacterioph	473	6	1.2	117	8	Q9XMS5	Q9XMS5 tetrahymena
401	6	1.2	95	11	P70644	P70644 rattus norv	474	6	1.2	118	5	019424	019424 caenorhabdi
402	6	1.2	95	12	Q37099	Q37099 human immun	475	6	1.2	118	12	092517	092517 beet virus
403	6	1.2	95	12	Q9WTL21	Q9WTL21 human immun	476	6	1.2	120	1	058193	058193 pyrococcus
404	6	1.2	95	12	Q9WTL16	Q9WTL16 human immun	477	6	1.2	121	2	049098	049098 mycoplasma
405	6	1.2	95	12	Q9WCS2	Q9WCS2 human immun	478	6	1.2	121	10	023089	023089 arabidopsis
406	6	1.2	95	12	Q9WCM4	Q9WCM4 human immun	479	6	1.2	123	12	Q9YWB1	Q9YWB1 human immun
407	6	1.2	95	12	Q9WCL1	Q9WCL1 human immun	480	6	1.2	124	2	068464	068464 rickettsia
408	6	1.2	95	12	Q9WCH9	Q9WCH9 human immun	481	6	1.2	124	2	068465	068465 rickettsia
409	6	1.2	95	12	Q9W9H1	Q9W9H1 human immun	482	6	1.2	125	2	Q9XD69	Q9XD69 legionella
410	6	1.2	95	13	Q73812	Q73812 morone saxa	483	6	1.2	125	12	P90116	P90116 human immun
411	6	1.2	96	12	Q9WCU1	Q9WCU1 human immun	484	6	1.2	127	2	Q9WYH4	Q9WYH4 thermotoga
412	6	1.2	96	12	Q9WCU0	Q9WCU0 human immun	485	6	1.2	128	2	P94214	P94214 lactobacill
413	6	1.2	96	12	Q9WCS7	Q9WCS7 human immun	486	6	1.2	128	3	059957	059957 agaricus bi
414	6	1.2	96	12	Q9WCP6	Q9WCP6 human immun	487	6	1.2	128	10	Q9XFL7	Q9XFL7 phaseolus v
415	6	1.2	96	12	Q9WCM3	Q9WCM3 human immun	488	6	1.2	131	2	P94342	P94342 burholderi
416	6	1.2	96	12	Q9WCM2	Q9WCM2 human immun	489	6	1.2	132	5	Q9XJ97	Q9XJ97 lampetra fl
417	6	1.2	96	12	Q9WCL8	Q9WCL8 human immun	490	6	1.2	133	2	Q9ZK84	Q9ZK84 helicobacte
418	6	1.2	96	12	Q9W4A3	Q9W4A3 human immun	491	6	1.2	133	12	Q9ZEP9	Q9ZEP9 pseudomonas
419	6	1.2	96	12	Q9W9V2	Q9W9V2 human immun	492	6	1.2	134	12	Q66118	Q66118 cucumber mo
420	6	1.2	98	2	Q31026	Q31026 vibrio chol	493	6	1.2	134	12	Q74951	Q74951 human immun
421	6	1.2	98	5	Q76496	Q76496 heliothis v	494	6	1.2	134	12	P88703	P88703 human immun
422	6	1.2	98	12	Q07958	Q07958 human immun	495	6	1.2	134	12	P90138	P90138 human immun
423	6	1.2	99	12	Q9WCH1	Q9WCH1 human immun	496	6	1.2	136	2	007663	007663 enterococcu
424	6	1.2	100	2	Q84034	Q84034 chlamydia t	497	6	1.2	136	5	045207	045207 anopheles g
425	6	1.2	103	2	Q49507	Q49507 mycoplasma	498	6	1.2	136	5	P90543	P90543 eupteris cr
426	6	1.2	103	12	Q87075	Q87075 pseudorabie	499	6	1.2	137	2	Q9ZIC8	Q9ZIC8 listeria mo
427	6	1.2	103	12	Q9WBF4	Q9WBF4 human immun	500	6	1.2	137	5	Q96334	Q96334 dirofilaria
428	6	1.2	104	12	Q9WBD8	Q9WBD8 human immun	501	6	1.2	137	5	Q96335	Q96335 brugia mala
429	6	1.2	105	2	Q85812	Q85812 streptococc	502	6	1.2	137	5	Q96336	Q96336 onchocerca
430	6	1.2	105	4	Q15944	Q15944 homo sapien	503	6	1.2	137	9	Q38374	Q38374 lactococcus
431	6	1.2	105	11	Q70155	Q70155 mus musculu	504	6	1.2	139	3	Q74900	Q74900 schistosach
432	6	1.2	106	12	Q9WCT6	Q9WCT6 human immun	505	6	1.2	139	5	P90945	P90945 caenorhabdi
433	6	1.2	107	12	Q90587	Q90587 human immun	506	6	1.2	139	12	Q75346	Q75346 human immun
434	6	1.2	107	12	Q90588	Q90588 human immun	507	6	1.2	139	12	Q9WAE5	Q9WAE5 influenza c
435	6	1.2	107	12	Q90590	Q90590 human immun	508	6	1.2	139	12	Q9WAD0	Q9WAD0 influenza c
436	6	1.2	107	12	Q9YRT3	Q9YRT3 human immun	509	6	1.2	139	12	Q9WAD8	Q9WAD8 influenza c
437	6	1.2	107	12	Q9YRT2	Q9YRT2 human immun	510	6	1.2	139	12	Q9WAD6	Q9WAD6 influenza c
438	6	1.2	107	12	Q9YRT0	Q9YRT0 human immun	511	6	1.2	139	12	Q9WAD4	Q9WAD4 influenza c
439	6	1.2	107	12	Q9YRS9	Q9YRS9 human immun	512	6	1.2	139	12	Q9WBT3	Q9WBT3 influenza c
440	6	1.2	107	12	Q9YRS8	Q9YRS8 human immun	513	6	1.2	139	12	Q9WBT5	Q9WBT5 influenza c
441	6	1.2	107	12	Q9YRS7	Q9YRS7 human immun	514	6	1.2	140	5	Q17821	Q17821 caenorhabdi
442	6	1.2	107	12	Q9YRS5	Q9YRS5 human immun	515	6	1.2	142	2	050753	050753 borrelia bu
443	6	1.2	107	12	Q9YRS4	Q9YRS4 human immun	516	6	1.2	142	5	Q97167	Q97167 culic pipie
444	6	1.2	107	12	Q9YRS3	Q9YRS3 human immun	517	6	1.2	142	6	Q9XSL4	Q9XSL4 sus scrofa
445	6	1.2	107	12	Q9YRS2	Q9YRS2 human immun	518	6	1.2	142	11	Q61522	Q61522 mus musculu
446	6	1.2	107	12	Q9YRP2	Q9YRP2 human immun	519	6	1.2	143	2	Q48729	Q48729 lactococcus
447	6	1.2	107	12	Q9YRX5	Q9YRX5 human immun	520	6	1.2	143	5	Q20291	Q20291 caenorhabdi
448	6	1.2	107	12	Q9WLD5	Q9WLD5 human immun	521	6	1.2	143	5	Q76223	Q76223 trypanosoma
449	6	1.2	108	5	Q99099	Q99099 onchocerca	522	6	1.2	143	9	Q9ZAX7	Q9ZAX7 bacterioph
450	6	1.2	108	12	Q55902	Q55902 human immun	523	6	1.2	144	5	Q01257	Q01257 caenorhabdi
451	6	1.2	109	1	Q58646	Q58646 pyrococcus	524	6	1.2	144	9	Q48447	Q48447 bacterioph
452	6	1.2	110	6	Q46627	Q46627 bos taurus	525	6	1.2	145	6	Q18926	Q18926 sus scrofa
453	6	1.2	110	12	Q96712	Q96712 cucumber mo	526	6	1.2	146	1	Q58599	Q58599 methanococc
454	6	1.2	111	2	Q50328	Q50328 escherichia	527	6	1.2	146	2	Q50021	Q50021 mycobacteri
455	6	1.2	111	12	Q39437	Q39437 cucumber mo	528	6	1.2	146	10	Q24024	Q24024 lycopersico
456	6	1.2	111	12	Q86784	Q86784 cucumber mo	529	6	1.2	146	10	Q41514	Q41514 tritpsacum a
457	6	1.2	111	12	Q9WA72	Q9WA72 cucumber mo	530	6	1.2	147	2	Q9WYK3	Q9WYK3 thermotoga

531	6	1.2	149	2	005444	005444 mycobacteri	604	1.2	191	5	026511	026511 schistosoma
532	6	1.2	149	2	09wxf4	09wxf4 thermotoga	605	1.2	191	5	003528	003528 schistosoma
533	6	1.2	150	4	09y4m1	09y4m1 homo sapien	606	1.2	191	5	001372	001372 schistosoma
534	6	1.2	150	5	002581	002581 inciliaria f	607	1.2	192	11	055109	055109 mus musculu
535	6	1.2	150	5	017962	017962 caenorhabdi	608	1.2	193	5	P90560	P90560 oxytricha f
536	6	1.2	150	12	098778	098778 vesicular s	609	1.2	193	5	094976	094976 neisseria f
537	6	1.2	150	12	098779	098779 vesicular s	610	1.2	194	2	069747	069747 streptococ
538	6	1.2	150	12	098780	098780 vesicular s	611	1.2	194	2	086479	086479 schistosoma
539	6	1.2	150	12	098790	098790 vesicular s	612	1.2	194	2	096853	096853 schistosoma
540	6	1.2	150	12	098791	098791 vesicular s	613	1.2	194	2	P28171	P28171 capra hircu
541	6	1.2	150	12	098787	098787 vesicular s	614	1.2	195	6	P75894	P75894 escherichia
542	6	1.2	150	12	098789	098789 vesicular s	615	1.2	197	1	059040	059040 pyrococcus
543	6	1.2	151	5	025283	025283 loligo fortb	616	1.2	197	10	Q41837	Q41837 zea mays (m
544	6	1.2	152	5	016955	016955 acanthamoeb	617	1.2	198	1	057724	057724 pyrococcus
545	6	1.2	152	5	025314	025314 leishmania	618	1.2	198	4	09y6k8	09y6k8 homo sapien
546	6	1.2	154	5	029686	029686 archaeoglob	619	1.2	199	2	087795	087795 prochlorococ
547	6	1.2	154	5	017739	017739 caenorhabdi	620	1.2	199	10	065687	065687 arabidopsis
548	6	1.2	155	10	P93828	P93828 arabidopsis	621	1.2	200	2	P73246	P73246 synecocyst
549	6	1.2	157	1	026871	026871 methanobact	622	1.2	200	12	092239	092239 human immun
550	6	1.2	157	2	047443	047443 escherichia	623	1.2	201	4	099935	099935 homo sapien
551	6	1.2	158	1	058766	058766 pyrococcus	624	1.2	201	12	Q9wkm9	Q9wkm9 human immun
552	6	1.2	158	2	P71105	P71105 clostridium	625	1.2	202	5	073378	073378 plasmodium
553	6	1.2	158	2	006277	006277 mycobacteri	626	1.2	202	10	Q41643	Q41643 volvox cart
554	6	1.2	158	10	081775	081775 arabidopsis	627	1.2	202	12	Q71827	Q71827 human immun
555	6	1.2	160	3	Q12540	Q12540 agaricus bl	628	1.2	202	12	Q9yrc7	Q9yrc7 human immun
556	6	1.2	161	2	09x6m4	09x6m4 streptococ	629	1.2	203	3	007451	007451 saccharomyc
557	6	1.2	162	2	033056	033056 mycobacteri	630	1.2	203	5	061828	061828 caenorhabdi
558	6	1.2	162	5	061829	061829 caenorhabdi	631	1.2	203	12	070726	070726 human immun
559	6	1.2	162	10	Q9xw56	Q9xw56 arabidopsis	632	1.2	204	3	008907	008907 saccharomyc
560	6	1.2	163	2	059307	059307 corynebacte	633	1.2	205	1	028634	028634 archaeoglob
561	6	1.2	166	2	030415	030415 lactococcus	634	1.2	205	12	Q9wll7	Q9wll7 human immun
562	6	1.2	167	2	067892	067892 aquifex aeo	635	1.2	206	5	061761	061761 caenorhabdi
563	6	1.2	168	1	057996	057996 pyrococcus	636	1.2	206	5	045868	045868 caenorhabdi
564	6	1.2	168	2	067861	067861 aquifex aeo	637	1.2	208	2	083947	083947 trepomema p
565	6	1.2	168	2	046662	046662 brucella ab	638	1.2	208	2	09x612	09x612 streptococ
566	6	1.2	168	5	Q9xay0	Q9xay0 pseudomonas	639	1.2	208	5	Q22156	Q22156 caenorhabdi
567	6	1.2	169	13	Q93343	Q93343 xenopus lae	640	1.2	208	8	078939	078939 agroclybe ae
568	6	1.2	170	2	Q9wv88	Q9wv88 pseudomonas	641	1.2	208	12	090025	090025 human immun
569	6	1.2	170	5	061171	061171 entodinium	642	1.2	209	2	Q9xc52	Q9xc52 legionella
570	6	1.2	171	5	046254	046254 scaptomyza	643	1.2	209	12	085953	085953 human immun
571	6	1.2	171	5	032418	032418 staphylococ	644	1.2	210	5	Q25448	Q25448 musca domes
572	6	1.2	172	2	P76572	P76572 escherichia	645	1.2	210	6	062813	062813 papio hamad
573	6	1.2	172	2	Q9z6n8	Q9z6n8 chlamydia p	646	1.2	210	8	078286	078286 palmaria pa
574	6	1.2	172	2	Q9x879	Q9x879 streptomyce	647	1.2	210	11	070377	070377 rattus norv
575	6	1.2	173	2	047446	047446 escherichia	648	1.2	210	11	035620	035620 mus musculu
576	6	1.2	173	2	047190	047190 escherichia	649	1.2	210	11	009044	009044 mus musculu
577	6	1.2	173	2	Q9xay5	Q9xay5 pseudomonas	650	1.2	211	2	026053	026053 helicobacte
578	6	1.2	173	2	Q9xay5	Q9xay5 pseudomonas	651	1.2	211	2	09zj99	09zj99 helicobacte
579	6	1.2	173	5	061172	061172 entodinium	652	1.2	211	2	09x2x4	09x2x4 bacillus an
580	6	1.2	175	1	028291	028291 archaeoglob	653	1.2	213	2	025918	025918 helicobacte
581	6	1.2	175	2	047449	047449 escherichia	654	1.2	213	2	P73135	P73135 synecocyst
582	6	1.2	175	2	003942	003942 bacterioph	655	1.2	214	2	Q9ydb2	Q9ydb2 streptomyce
583	6	1.2	175	2	Q9z6n8	Q9z6n8 helicobacte	656	1.2	215	5	Q25445	Q25445 musca domes
584	6	1.2	177	2	Q9z6n8	Q9z6n8 helicobacte	657	1.2	215	5	Q25450	Q25450 musca domes
585	6	1.2	177	11	054841	054841 mus musculu	658	1.2	215	12	Q9yP67	Q9yP67 sweet potat
586	6	1.2	181	1	028254	028254 archaeoglob	659	1.2	215	5	Q9xw55	Q9xw55 caenorhabdi
587	6	1.2	183	5	018248	018248 caenorhabdi	660	1.2	216	5	P90053	P90053 human immun
588	6	1.2	184	5	022026	022026 caenorhabdi	661	1.2	216	12	Q9xw50	Q9xw50 caenorhabdi
589	6	1.2	184	10	080364	080364 arabidopsis	662	1.2	218	5	Q9xw50	Q9xw50 caenorhabdi
590	6	1.2	184	12	088133	088133 visna virus	663	1.2	218	11	064401	064401 cricetus
591	6	1.2	185	13	P79857	P79857 pleurodeles	664	1.2	218	11	060466	060466 cricetus
592	6	1.2	186	2	Q9wyt4	Q9wyt4 thermotoga	665	1.2	219	5	025514	025514 helicobacte
593	6	1.2	186	5	001589	001589 caenorhabdi	666	1.2	219	5	017346	017346 caenorhabdi
594	6	1.2	186	10	Q9xw58	Q9xw58 arabidopsis	667	1.2	220	2	069465	069465 mycobacteri
595	6	1.2	187	12	Q9yrm8	Q9yrm8 human immun	668	1.2	221	2	Q9w2x2	Q9w2x2 bacillus su
596	6	1.2	187	12	Q9yrc6	Q9yrc6 human immun	669	1.2	221	10	Q96477	Q96477 lycopersico
597	6	1.2	188	3	Q99321	Q99321 saccharomyc	670	1.2	222	2	086632	086632 streptomyce
598	6	1.2	189	5	023631	023631 cryptospori	671	1.2	222	2	09x622	09x622 bacillus sp
599	6	1.2	190	1	029539	029539 archaeoglob	672	1.2	222	12	082125	082125 viral hemor
600	6	1.2	190	4	046236	046236 drosophila	673	1.2	222	12	008475	008475 viral hemor
601	6	1.2	191	4	099491	099491 homo sapien	674	1.2	223	10	Q41429	Q41429 solanum tub
602	6	1.2	191	5	044372	044372 schistosoma	675	1.2	223	12	Q76769	Q76769 human immun
603	6	1.2	191	5	017091	017091 caenorhabdi	676	1.2	223	12	Q76770	Q76770 human immun

677	6	1.2	223	076771	076771 human immun	750	6	1.2	256	6	002756	002756 bos taurus
678	6	1.2	223	076772	076772 human immun	751	6	1.2	257	2	031204	031204 pasteurella
679	6	1.2	224	066646	066646 aquifex aeo	752	6	1.2	257	2	084195	084195 chlamydia t
680	6	1.2	225	092571	092571 zymomonas m	753	6	1.2	257	10	09XG32	09XG32 guillamdia t
681	6	1.2	225	092033	092033 arabidopsis	754	6	1.2	258	1	029264	029264 archaeoglob
682	6	1.2	226	026341	026341 methanobact	755	6	1.2	258	1	09YDB9	09YDB9 aeropyrum p
683	6	1.2	226	006319	006319 mycobacteri	756	6	1.2	258	2	056836	056836 amoeba prot
684	6	1.2	226	083590	083590 treponema p	757	6	1.2	258	2	09X414	09X414 chlamydiophl
685	6	1.2	226	024082	024082 diosiphila	758	6	1.2	258	10	064595	064595 arabidopsis
686	6	1.2	226	098686	098686 simian cyto	759	6	1.2	259	10	082798	082798 arabidopsis
687	6	1.2	227	09WM59	09WM59 pseudomonas	760	6	1.2	259	12	096867	096867 simian cyto
688	6	1.2	227	09X107	09X107 falco pereg	761	6	1.2	260	2	025113	025113 helicobacte
689	6	1.2	230	085466	085466 bacillus ce	762	6	1.2	260	2	09XAB5	09XAB5 bacillus ce
690	6	1.2	230	017503	017503 caenorhabdi	763	6	1.2	262	2	09WZP6	09WZP6 thermotoga
691	6	1.2	231	095372	095372 homo sapien	764	6	1.2	262	5	097313	097313 plasmidum
692	6	1.2	231	092WS9	092WS9 arabidopsis	765	6	1.2	262	12	092547	092547 epizootic h
693	6	1.2	231	09WTL7	09WTL7 mus musculu	766	6	1.2	262	14	009195	009195 unidentified
694	6	1.2	232	053668	053668 streptomyce	767	6	1.2	263	12	084599	084599 chlamydia t
695	6	1.2	232	055942	055942 synechocyst	768	6	1.2	263	12	071100	071100 bovine aden
696	6	1.2	233	074284	074284 synechocyst	769	6	1.2	263	12	09WDX5	09WDX5 human immun
697	6	1.2	233	096361	096361 hyphantria	770	6	1.2	265	2	007791	007791 mycobacteri
698	6	1.2	234	056180	056180 synechococ	771	6	1.2	265	5	016807	016807 hirudo medi
699	6	1.2	235	0928L9	0928L9 chlamydia p	772	6	1.2	265	11	054830	054830 mus musculu
700	6	1.2	235	002442	002442 globodera r	773	6	1.2	266	2	09XG32	09XG32 streptomyce
701	6	1.2	235	098224	098224 chlamydomon	774	6	1.2	266	11	035334	035334 mus musculu
702	6	1.2	235	033368	033368 chlamydomon	775	6	1.2	267	8	048006	048006 thomomys ma
703	6	1.2	236	028882	028882 archaeoglob	776	6	1.2	267	8	035178	035178 neacoms sp
704	6	1.2	237	09X830	09X830 streptomyce	777	6	1.2	268	4	075806	075806 homo sapien
705	6	1.2	238	090337	090337 saquaro cac	778	6	1.2	268	14	009194	009194 unidentified
706	6	1.2	239	029017	029017 archaeoglob	779	6	1.2	269	10	039360	039360 brassica na
707	6	1.2	239	09YAS7	09YAS7 aeropyrum p	780	6	1.2	270	2	050653	050653 paracoccus
708	6	1.2	239	090778	090778 caenorhabdi	781	6	1.2	270	2	073854	073854 synechocyst
709	6	1.2	240	096178	096178 plasmidum	782	6	1.2	270	3	094105	094105 pleurotus d
710	6	1.2	241	032827	032827 legionella	783	6	1.2	270	10	024141	024141 nicotiana t
711	6	1.2	241	084728	084728 chlamydia t	784	6	1.2	271	2	005879	005879 mycobacteri
712	6	1.2	241	022849	022849 arabidopsis	785	6	1.2	271	10	004633	004633 arabidopsis
713	6	1.2	242	032767	032767 legionella	786	6	1.2	272	10	022806	022806 arabidopsis
714	6	1.2	242	018238	018238 caenorhabdi	787	6	1.2	273	10	096740	096740 arabidopsis
715	6	1.2	242	096136	096136 plasmidum	788	6	1.2	274	10	081395	081395 atrichrium
716	6	1.2	242	028075	028075 bos taurus	789	6	1.2	274	12	089048	089048 vesicular s
717	6	1.2	243	057997	057997 pyrococcus	790	6	1.2	275	10	064700	064700 arabidopsis
718	6	1.2	243	092FU0	092FU0 streptococ	791	6	1.2	277	1	028441	028441 archaeoglob
719	6	1.2	243	080208	080208 methanobact	792	6	1.2	277	12	09WEQ0	09WEQ0 human immun
720	6	1.2	244	095601	095601 rhizobium e	793	6	1.2	277	12	09WDX7	09WDX7 human immun
721	6	1.2	244	016421	016421 caenorhabdi	794	6	1.2	278	2	053104	053104 rhodobacter
722	6	1.2	245	018039	018039 caenorhabdi	795	6	1.2	278	2	092811	092811 chlamydia p
723	6	1.2	246	090811	090811 gallus gall	796	6	1.2	278	5	061831	061831 caenorhabdi
724	6	1.2	247	053566	053566 haemophilus	797	6	1.2	279	1	09YBPI	09YBPI aeropyrum p
725	6	1.2	247	020284	020284 caenorhabdi	798	6	1.2	279	12	09WEQ6	09WEQ6 human immun
726	6	1.2	247	092914	092914 actinidia c	799	6	1.2	280	2	09WYGI	09WYGI thermotoga
727	6	1.2	247	070140	070140 mus musculu	800	6	1.2	280	10	041602	041602 tritillum du
728	6	1.2	248	043033	043033 schizosacch	801	6	1.2	281	2	054329	054329 streptomyce
729	6	1.2	248	045450	045450 caenorhabdi	802	6	1.2	281	6	077728	077728 ovis aries
730	6	1.2	248	09WTF6	09WTF6 human immun	803	6	1.2	281	8	09XNF7	09XNF7 elephas max
731	6	1.2	249	046357	046357 caenorhabdi	804	6	1.2	282	12	057164	057164 spinach lat
732	6	1.2	250	09Z1A2	09Z1A2 clostridium	805	6	1.2	283	2	085831	085831 sphingomona
733	6	1.2	250	0927V0	0927V0 chlamydia p	806	6	1.2	284	2	0928Y3	0928Y3 chlamydia p
734	6	1.2	250	090779	090779 caenorhabdi	807	6	1.2	284	4	014874	014874 homo sapien
735	6	1.2	250	017721	017721 caenorhabdi	808	6	1.2	284	4	09XMF9	09XMF9 elephas max
736	6	1.2	251	032768	032768 legionella	809	6	1.2	285	1	09YAB5	09YAB5 aeropyrum p
737	6	1.2	252	051905	051905 proteus mir	810	6	1.2	285	2	074637	074637 synechocyst
738	6	1.2	252	094663	094663 schizosacch	811	6	1.2	285	5	094553	094553 dugesia tlg
739	6	1.2	252	067894	067894 arabidopsis	812	6	1.2	285	5	045492	045492 caenorhabdi
740	6	1.2	253	083296	083296 aquifex aeo	813	6	1.2	285	8	020957	020957 cryphonectr
741	6	1.2	253	083296	083296 treponema p	814	6	1.2	285	8	09XMF8	09XMF8 elephas max
742	6	1.2	253	026869	026869 typanosoma	815	6	1.2	285	10	082166	082166 arabidopsis
743	6	1.2	253	09WY65	09WY65 rattus norv	816	6	1.2	286	2	024958	024958 helicobacte
744	6	1.2	255	054169	054169 streptomyce	817	6	1.2	286	12	088977	088977 venezuelan
745	6	1.2	255	09WZ47	09WZ47 thermotoga	818	6	1.2	287	10	048781	048781 arabidopsis
746	6	1.2	256	026335	026335 methanobact	819	6	1.2	288	3	094366	094366 schizosacch
747	6	1.2	256	031618	031618 bacillus su	820	6	1.2	289	1	026566	026566 methanobact
748	6	1.2	256	095844	095844 staphylococ	821	6	1.2	289	2	045377	045377 caulobacter
749	6	1.2	256	007419	007419 typanosoma	822	6	1.2	289	2	067230	067230 aquifex aeo

823	6	1.2	289	11	062204	062204 mus musculu	896	6	1.2	320	10	Q41324	Q41324 stylosanthe
824	6	1.2	292	5	09XZL8	09XZL8 drosophila	897	6	1.2	321	2	09WYL7	09WYL7 thermotoga
825	6	1.2	293	2	Q46564	Q46564 bacteroides	898	6	1.2	321	10	09XIV9	09XIV9 nicotiana t
826	6	1.2	293	2	Q52694	Q52694 roseburia c	899	6	1.2	321	11	054851	054851 mus musculu
827	6	1.2	293	3	09XBUD	09XBUD porphyromon	900	6	1.2	321	11	070610	070610 rattus norv
828	6	1.2	293	5	016423	016423 caenorhabd1	901	6	1.2	321	11	070610	070610 rattus norv
829	6	1.2	294	2	009257	009257 synechococc	902	6	1.2	322	2	Q46003	Q46003 campylobact
830	6	1.2	295	2	09WMR5	09WMR5 pseudomonas	903	6	1.2	323	4	075028	075028 homo sapien
831	6	1.2	295	4	099322	099322 homo sapien	904	6	1.2	323	4	09Y698	09Y698 homo sapien
832	6	1.2	296	1	09YER8	09YER8 aeropyrum p	905	6	1.2	323	5	066835	066835 drosophila
833	6	1.2	296	2	053053	053053 listeria iv	906	6	1.2	323	11	088602	088602 mus musculu
834	6	1.2	296	5	017002	017002 caenorhabd1	907	6	1.2	324	1	029940	029940 archaeoglob
835	6	1.2	296	6	09ZT04	09ZT04 dicostonyx	908	6	1.2	324	10	Q41097	Q41097 populus tri
836	6	1.2	297	2	054383	054383 rickettsia	909	6	1.2	325	2	Q55392	Q55392 synechocyst
837	6	1.2	297	2	P71451	P71451 listeria mo	910	6	1.2	325	13	Q42169	Q42169 brachydanio
838	6	1.2	297	9	080085	080085 bacterioph	911	6	1.2	326	2	P76071	P76071 escherichia
839	6	1.2	297	10	Q41122	Q41122 phaeococcus v	912	6	1.2	326	2	086322	086322 mycobacteri
840	6	1.2	298	1	059546	059546 pyrococcus	913	6	1.2	326	4	012251	012251 saccharomyc
841	6	1.2	298	2	P71702	P71702 mycobacteri	914	6	1.2	326	4	000272	000272 homo sapien
842	6	1.2	298	8	Q9XKX9	Q9XKX9 agrocyste ae	915	6	1.2	326	8	099019	099019 solanum dem
843	6	1.2	299	5	Q22856	Q22856 caenorhabd1	916	6	1.2	326	9	Q38213	Q38213 bacterioph
844	6	1.2	300	2	069782	069782 stinorhizob1	917	6	1.2	326	10	Q24392	Q24392 solanum tub
845	6	1.2	301	2	P94627	P94627 clostridium	918	6	1.2	326	10	055846	055846 methicococ
846	6	1.2	301	2	Q44404	Q44404 agrobacteri	919	6	1.2	327	2	069447	069447 methicococ
847	6	1.2	301	2	Q47396	Q47396 escherichia	920	6	1.2	329	5	018461	018461 haematobia
848	6	1.2	301	2	Q53756	Q53756 streptomyce	921	6	1.2	329	5	018463	018463 haematobia
849	6	1.2	301	5	09Y201	09Y201 ephydaria f	922	6	1.2	329	12	066100	066100 carnationa f
850	6	1.2	301	8	09ZV07	09ZV07 dicostonyx	923	6	1.2	329	12	09YMG5	09YMG5 lymantria d
851	6	1.2	302	5	018318	018318 caenorhabd1	924	6	1.2	330	1	058121	058121 pyrococcus
852	6	1.2	303	2	Q9ZG55	Q9ZG55 leptospira	925	6	1.2	330	2	Q9XSV7	Q9XSV7 streptomyce
853	6	1.2	303	8	Q9ZT05	Q9ZT05 dicostonyx	926	6	1.2	332	2	062627	062627 rattus norv
854	6	1.2	304	2	066991	066991 aquifex aeo	927	6	1.2	332	11	062627	062627 rattus norv
855	6	1.2	305	1	Q26838	Q26838 methanobact	928	6	1.2	333	2	09MYW8	09MYW8 thermotoga
856	6	1.2	305	2	067486	067486 aquifex aeo	929	6	1.2	334	2	Q25836	Q25836 helicobacte
857	6	1.2	305	2	007166	007166 mycobacteri	930	6	1.2	334	2	Q9ZVY8	Q9ZVY8 helicobacte
858	6	1.2	305	8	Q9ZV06	Q9ZV06 dicostonyx	931	6	1.2	334	3	014075	014075 schizosacch
859	6	1.2	305	13	Q9I346	Q9I346 xenopus lae	932	6	1.2	334	12	072733	072733 cowpox viru
860	6	1.2	306	2	Q47696	Q47696 escherichia	933	6	1.2	335	8	Q9ZV31	Q9ZV31 elephantiu
861	6	1.2	306	10	065623	065623 arabidopsis	934	6	1.2	336	1	030070	030070 archaeoglob
862	6	1.2	306	10	081845	081845 arabidopsis	935	6	1.2	336	2	Q49675	Q49675 mycobacteri
863	6	1.2	307	2	Q9ZT74	Q9ZT74 chlamydia p	936	6	1.2	336	2	P77749	P77749 mycobacteri
864	6	1.2	307	5	077394	077394 plasmodium	937	6	1.2	337	8	078674	078674 porphyra dl
865	6	1.2	308	2	P72053	P72053 mycobacteri	938	6	1.2	337	4	000556	000556 homo sapien
866	6	1.2	308	2	067168	067168 aquifex aeo	939	6	1.2	337	5	077122	077122 lacosoma ch
867	6	1.2	308	5	Q45310	Q45310 caenorhabd1	940	6	1.2	338	2	Q45254	Q45254 bradyrhizob
868	6	1.2	309	5	Q23138	Q23138 caenorhabd1	941	6	1.2	338	2	P73887	P73887 synechocyst
869	6	1.2	309	10	P93550	P93550 spinacia ol	942	6	1.2	339	2	P94915	P94915 mycobacteri
870	6	1.2	310	2	052737	052737 streptococc	943	6	1.2	339	2	Q9XD85	Q9XD85 corynebacte
871	6	1.2	310	2	052738	052738 streptococc	944	6	1.2	340	1	09X7U9	09X7U9 streptomyce
872	6	1.2	310	2	052739	052739 streptococc	945	6	1.2	341	1	Q9YAF1	Q9YAF1 aeropyrum p
873	6	1.2	310	2	052740	052740 streptococc	946	6	1.2	341	5	095565	095565 homo sapien
874	6	1.2	310	2	052741	052741 streptococc	947	6	1.2	341	5	097103	097103 crithidia l
875	6	1.2	310	2	052742	052742 streptococc	948	6	1.2	341	5	097106	097106 trypanosoma
876	6	1.2	310	2	052743	052743 streptococc	949	6	1.2	341	5	097107	097107 trypanosoma
877	6	1.2	310	2	Q9ZBE7	Q9ZBE7 mycobacteri	950	6	1.2	342	1	058832	058832 pyrococcus
878	6	1.2	310	12	Q89181	Q89181 vaccinia vl	951	6	1.2	342	4	075796	075796 escherichia
879	6	1.2	312	2	Q25095	Q25095 helicobacte	952	6	1.2	342	4	075796	075796 homo sapien
880	6	1.2	312	4	095566	095566 homo sapien	953	6	1.2	342	8	098606	098606 porphyra le
881	6	1.2	312	4	016417	016417 caenorhabd1	954	6	1.2	342	12	055831	055831 stratford v
882	6	1.2	313	8	048005	048005 thomomys ma	955	6	1.2	344	2	P94154	P94154 alcaligenes
883	6	1.2	314	2	033175	033175 mycobacteri	956	6	1.2	344	2	067238	067238 pseudomonas
884	6	1.2	314	3	059909	059909 aspergillus	957	6	1.2	344	2	Q9Z9H2	Q9Z9H2 pseudomonas
885	6	1.2	315	10	P93549	P93549 spinacia ol	958	6	1.2	345	5	017809	017809 caenorhabd1
886	6	1.2	315	10	Q24336	Q24336 raphanus sa	959	6	1.2	345	2	P75340	P75340 mycoplasma
887	6	1.2	316	5	018460	018460 haematobia	960	6	1.2	345	2	030883	030883 borrelia bu
888	6	1.2	316	5	018460	018462 haematobia	961	6	1.2	345	6	P73981	P73981 synechocyst
889	6	1.2	317	1	029954	029954 archaeoglob	962	6	1.2	345	6	Q27951	Q27951 bos taurus
890	6	1.2	317	2	066374	066374 acetobacter	963	6	1.2	346	8	078890	078890 anolis marc
891	6	1.2	317	2	066538	066538 aquifex aeo	964	6	1.2	346	5	Q25871	Q25871 helicobacte
892	6	1.2	318	1	057922	057922 pyrococcus	965	6	1.2	346	5	P91677	P91677 drosophila
893	6	1.2	319	2	006448	006448 serratia ma	966	6	1.2	347	8	019859	019859 porphyra am
894	6	1.2	320	2	068457	068457 acinetobact	967	6	1.2	347	8	P95603	P95603 alcaligenes
895	6	1.2	320	5	Q93616	Q93616 caenorhabd1	968	6	1.2	347	3	Q02428	Q02428 saccharomyc

```

969 6 1.2 347 3 012452 saccharomy
970 6 1.2 347 8 033165 saccharomy
971 6 1.2 347 8 033164 saccharomy
972 6 1.2 347 8 033087 saccharomy
973 6 1.2 347 8 033071 saccharomy
974 6 1.2 347 8 032634 saccharomy
975 6 1.2 347 8 032590 saccharomy
976 6 1.2 347 8 032573 saccharomy
977 6 1.2 347 8 032557 saccharomy
978 6 1.2 348 2 083091 saccharomy
979 6 1.2 348 5 024084 saccharomy
980 6 1.2 348 5 024084 saccharomy
981 6 1.2 348 5 024726 saccharomy
982 6 1.2 349 1 028654 saccharomy
983 6 1.2 349 4 093355 saccharomy
984 6 1.2 349 4 019858 saccharomy
985 6 1.2 349 10 038705 saccharomy
986 6 1.2 349 11 092025 saccharomy
987 6 1.2 349 12 071072 saccharomy
988 6 1.2 350 2 033543 saccharomy
989 6 1.2 350 2 047720 saccharomy
990 6 1.2 350 3 093758 saccharomy
991 6 1.2 350 5 027557 saccharomy
992 6 1.2 350 10 043070 saccharomy
993 6 1.2 350 12 071054 saccharomy
994 6 1.2 350 12 071055 saccharomy
995 6 1.2 350 12 071062 saccharomy
996 6 1.2 350 12 071063 saccharomy
997 6 1.2 350 12 071065 saccharomy
998 6 1.2 350 12 071070 saccharomy
999 6 1.2 350 12 071052 saccharomy
1000 6 1.2 350 12 071056 saccharomy

```

ALIGNMENTS

```

RESULT 1
ID 068697 PRELIMINARY; PRT; 326 AA.
AC 068697;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
DE V ANTIGEN.
GN ICRV.
OS Yersinia pestis.
OC Plasmid pCDL.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM;
RA HU P., ELLIOTT J., MCCREARY P., SKOMRONSKI E., GARNES J.,
RA KOHAYASHI A., CARRANO A.V., BRUBAKER R., GARCIA E.;
RA Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM;
RA MEDLINE; 98427122.
RA PERRY R.D., STRALEY S.C., FETHERSTON J.D., ROSE D.J., GREGOR J.,
RA BLATTNER F.R.;
RT "DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of
RT Yersinia pestis KIM5."
RL Infect. Immun. 66:4611-4623(1998).
DR EMBL; AF053946; AAC62574.1; -
DR EMBL; AF074612; AAC69799.1; -
KW Plasmid.
SQ SEQUENCE 326 AA; 37240 MW; 443BEDDC CRC32;

```

Query Match 46.6%; Score 243; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 7.1e-237;

```

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 ILKGGHYDNOQNGIKRYKEFLSSPTOMELRAFMVHPSLTADRIDDDILKVIYDSM 338
DB 84 ILKGGHYDNOQNGIKRYKEFLSSPTOMELRAFMVHPSLTADRIDDDILKVIYDSM 143
QY 339 NHHGDANSKLEELAEELAEIKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 398
DB 144 NHHGDANSKLEELAEELAEIKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 203
QY 399 EIRKASAEVILKMPOTTIOVDGSEKKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 456
DB 204 EIRKASAEVILKMPOTTIOVDGSEKKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 263
QY 459 NELSHFATTCSDKSRPLNDLVSOQKTQLOSDTSRNSAIEALNFIOKYDSVMORLDDT 518
DB 264 NELSHFATTCSDKSRPLNDLVSOQKTQLOSDTSRNSAIEALNFIOKYDSVMORLDDT 323
QY 519 SGK 521
DB 324 SGK 326

```

```

RESULT 2
ID 056853 PRELIMINARY; PRT; 334 AA.
AC 056853;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE V ANTIGEN (FRAGMENT).
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WA-314;
RA MEDLINE; 97162308.
RA ROEGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESEWANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by anti-
RT recombinant V antigen is dependent on polymorphism of V antigen."
RL Infect. Immun. 65:446-451(1997).
DR EMBL; X66796; CAA65588.1; -
FT NON_TER 334
SQ SEQUENCE 334 AA; 38241 MW; 71229360 CRC32;

```

Query Match 20.0%; Score 104; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.5e-96;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 279 ILKGGHYDNOQNGIKRYKEFLSSPTOMELRAFMVHPSLTADRIDDDILKVIYDSM 338
DB 84 ILKGGHYDNOQNGIKRYKEFLSSPTOMELRAFMVHPSLTADRIDDDILKVIYDSM 143
QY 339 NHHGDANSKLEELAEELAEIKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 398
DB 144 NHHGDANSKLEELAEELAEIKIYVIAQAEINKHLSSSGTINIHDKSINLMDKNLYGYTD 203
RESULT 3
ID 056896 PRELIMINARY; PRT; 334 AA.
AC 056896;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE V ANTIGEN (FRAGMENT).
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
RN [1]
RP SEQUENCE FROM N.A.

```


RC STRAIN-NCTC;
 MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESESMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96797; CAA65589.1; -.
 FT NON_TER 334 334
 SO SEQUENCE 334 AA; 38227 MW; F23FED3 CRC32;

Query Match 20.4%; Score 104; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 1.5e-96;
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 ILKGHYDNLONGIRKRVKEFLSSPTQWELRAFMVHFSLTADRIDDDILKVIYDSM 338
 DB 84 ILKGHYDNLONGIRKRVKEFLSSPTQWELRAFMVHFSLTADRIDDDILKVIYDSM 143
 OY 339 NHHGDARSKLREELAEITAEIKIYSVIOAEINKHLS 382
 DB 144 NHHGDARSKLREELAEITAEIKIYSVIOAEINKHLS 187

RESULT 4
 ID 056846 PRELIMINARY; PRT: 324 AA.
 AC 056846;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE V ANTIGEN.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y-108-P;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESESMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96799; CAA65591.1; -.
 FT NON_TER 324 324
 SO SEQUENCE 324 AA; 37194 MW; 5BD357FF CRC32;

Query Match 18.4%; Score 96; DB 2; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.8e-88;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 ILKGHYDNLONGIRKRVKEFLSSPTQWELRAFMVHFSLTADRIDDDILKVIYDSM 338
 DB 84 ILKGHYDNLONGIRKRVKEFLSSPTQWELRAFMVHFSLTADRIDDDILKVIYDSM 143
 OY 339 NHHGDARSKLREELAEITAEIKIYSVIOAEINKHLS 374
 DB 144 NHHGDARSKLREELAEITAEIKIYSVIOAEINKHLS 179

RESULT 5
 ID 056879 PRELIMINARY; PRT: 324 AA.
 AC 056879;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE V ANTIGEN.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESESMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96800; CAA65592.1; -.
 FT NON_TER 324 324
 SO SEQUENCE 324 AA; 37195 MW; E1CB300F CRC32;

Query Match 18.4%; Score 96; DB 2; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.8e-88;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 ILKGHYDNLONGIRKRVKEFLSSPTQWELRAFMVHFSLTADRIDDDILKVIYDSM 338
 DB 84 ILKGHYDNLONGIRKRVKEFLSSPTQWELRAFMVHFSLTADRIDDDILKVIYDSM 143
 OY 339 NHHGDARSKLREELAEITAEIKIYSVIOAEINKHLS 374
 DB 144 NHHGDARSKLREELAEITAEIKIYSVIOAEINKHLS 179

RESULT 6
 ID 056880 PRELIMINARY; PRT: 325 AA.
 AC 056880;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE V ANTIGEN (FRAGMENT).
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-8081;
 RX MEDLINE: 97162308.
 RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESESMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 recombinant V antigen is dependent on polymorphism of V antigen.";
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96798; CAA65590.1; -.
 FT NON_TER 325 325
 SO SEQUENCE 325 AA; 37260 MW; D6E88446 CRC32;

Query Match 17.9%; Score 93; DB 2; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.9e-85;
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 ILKGHYDNLONGIRKRVKEFLSSPTQWELRAFMVHFSLTADRIDDDILKVIYDSM 338
 DB 84 ILKGHYDNLONGIRKRVKEFLSSPTQWELRAFMVHFSLTADRIDDDILKVIYDSM 143
 OY 339 NHHGDARSKLREELAEITAEIKIYSVIOAEINK 371
 DB 144 NHHGDARSKLREELAEITAEIKIYSVIOAEINK 176

RESULT 7
 ID 087495 PRELIMINARY; PRT: 324 AA.
 AC 087495;
 DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE ICRV PROTEIN.
 GN ICRV.
 OS Yersinia enterocolitica.
 OG Plasmid pIV, and Plasmid pYve227.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-W22703;
 RA IIRIARIE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
 RT "Complete sequence of the Yersinia enterocolitica pYV plasmid."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-W22703;
 RA IIRIARIE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
 RT "Detailed genetic map of the pYVE227 plasmid of Yersinia
 enterocolitica serotype O:9."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF102990; AAD16815.1;
 KM plasmid.
 SQ SEQUENCE 324 AA; 37296 MW; 7C9A652F CRC32;

Query Match 13.1%; Score 68; DB 2; Length 324;
 Best Local Similarity 100.0%; Pred. No. 3.2e-60;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 SYSYKDNNELSHFATTCSDKSRPLNDLVSOQTQLSDITSFNSAIALNFIQKIDSV 510
 DB 256 SYSYKDNNELSHFATTCSDKSRPLNDLVSOQTQLSDITSFNSAIALNFIQKIDSV 315
 QY 511 MQRLLDDT 518
 DB 316 MQRLLDDT 323

RESULT 8
 056882
 ID 056882 PRELIMINARY; PRT; 324 AA.
 AC 056882;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE Y ANTIGEN.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y-96-P;
 RX MEDLINE; 97162308.
 RA ROOGERKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HERSEMANN J.;
 RT "Passive immunity to infection with Yersinia spp. mediated by anti-
 RT recombinant Y antigen is dependent on polymorphism of Y antigen."
 RL Infect. Immun. 65:446-451(1997).
 DR EMBL: X96801; CAA65593.1;
 SQ SEQUENCE 324 AA; 37162 MW; 56842DOA CRC32;

Query Match 12.9%; Score 67; DB 2; Length 324;
 Best Local Similarity 100.0%; Pred. No. 3.3e-59;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 279 ILKGGHYDNLQNGIKRVEKLESSPNTQWELRAFMAVHESLFDRIIDDILKIYVDSM 338
 DB 84 ILKGGHYDNLQNGIKRVEKLESSPNTQWELRAFMAVHESLFDRIIDDILKIYVDSM 143
 QY 339 NHGDAR 345
 DB 144 NHGDAR 150

RESULT 9
 016256
 ID 016256 PRELIMINARY; PRT; 168 AA.
 AC 016256;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)

DE WT1-WILMS' TUMOR SUPPRESSOR PROTEIN.
 GN WT1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 95166649.
 RA HAMILTON T.B., BARIJLA K.C., ROMANIUK P.J.;
 RT "High affinity binding sites for the Wilms' tumour suppressor protein
 RT WT1".
 RL Nucleic Acids Res. 23:277-284(1995).
 DR EMBL: S75264; AAB33443.1;
 DR HSSP; P08046; IALL.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
 DR PFM; PF00096; zf-C2H2; 4.
 DR PRINTS; PR00048; ZINC_FINGER.
 KW Zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 168 AA; 20165 MW; E86026CC CRC32;

Query Match 3.3%; Score 17; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 5.2e-09;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGHRRHHHHHSSGHI 17
 DB 1 MGHRRHHHHHSSGHI 17

RESULT 10
 030527
 ID 030527 PRELIMINARY; PRT; 294 AA.
 AC 030527;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE PCRV.
 GN PCRV.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
 OC Pseudomonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-388;
 RX MEDLINE; 98037517.
 RA YAHN T.L., MENDE-MUELLER L.M., FRIESE M.B., FRANK D.W.;
 RT "Identification of type III secreted products of the Pseudomonas
 RT aeruginosa exoenzyme S regulon."
 RL J. Bacteriol. 179:7165-7168(1997).
 DR EMBL: AF010149; AAC45935.1;
 SQ SEQUENCE 294 AA; 32283 MW; 0FAD2082 CRC32;

Query Match 2.5%; Score 13; DB 2; Length 294;
 Best Local Similarity 100.0%; Pred. No. 9.8e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 498 EALNRFIOKIDSV 510
 DB 274 EALNRFIOKIDSV 286

RESULT 11
 042186
 ID 042186 PRELIMINARY; PRT; 516 AA.
 AC 042186;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
 DE ZEPFL.
 GN Brachydanio rerio (Zebrafish) (Zebra danio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinidae; Rasbora; Danio.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 97322094.
 RA LUD D., LE DREAN Y., EKKER M., XIONG F., HEW C.L.;
 RT "Teleost FT2-F1 homolog and its splicing variant determine the
 RT expression of the salmon gonadotropin IIBeta subunit gene.";
 RL Mol. Endocrinol. 11:877-890(1997).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF014926; AAC60274.1; -.
 DR HSSP; P19793; 2ML.
 DR ZFIN; ZDB-GENE-990415-79; zff1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR PFAM; PF00104; hormone_rec; 1.
 DR PFAM; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STROIDFINGER.
 DR KEGG; K01111; Transcription regulation; DNA-binding; Nuclear protein;
 KM Zinc-finger.
 SQ SEQUENCE 516 AA; 58805 MW; 11AB2783 CRC32;

Query Match 2.38; Score 12; DB 13; Length 516;
 Best Local Similarity 100.0%; Pred. No. 0.0017;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HHHHHHHHSS 14
 |||||
 DB 189 HHHHHHHHSS 200

RESULT 12
 ID 075263 PRELIMINARY; PRT; 230 AA.
 AC 075263;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
 DE WIGSC.H.DJ0620P06.1 PROTEIN (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Euteleostei; Primates; Catarrhini; Hominiidae; Homo.
 RN [1]
 RA ANDREWS S., LACY M., DOBBELDE C.;
 RT "The sequence of Homo sapiens PAC clone DJ0620P06.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA WATERSTON R.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; ACC05550; AAC33152.1; -.
 DR PFAM; PF00046; homeobox; 1.
 RT NON_TER 230
 FT SEQUENCE 230 AA; 25476 MW; 1A804F2B CRC32;

Query Match 2.18; Score 11; DB 4; Length 230;
 Best Local Similarity 100.0%; Pred. No. 0.0081;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 DB 67 GHHHHHHHHH 77

RESULT 13
 ID 042980 PRELIMINARY; PRT; 420 AA.
 AC 042980;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
 DE HYPOTHETICAL 47.3 KD PROTEIN.
 GN SPBC17D11.01 OR SPBC20F10.11.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972H-;
 RA BECK A., REINHARDT R., WOOD V., RAJANDREAM M.A., BARRELL B.G.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-281 FROM N.A.
 RC STRAIN-972H-;
 RA LYNE M., MCDONALD R., RAJANDREAM M.A., BARRELL B.G., BECK A.,
 RA REINHARDT R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL031739; CAA21072.1; -.
 DR EMBL; AL021747; CAA16851.1; -.
 KM Hypothetical protein.
 SQ SEQUENCE 420 AA; 47257 MW; 35DC88FC CRC32;

Query Match 2.18; Score 11; DB 3; Length 420;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 DB 298 GHHHHHHHHH 308

RESULT 14
 ID P78807 PRELIMINARY; PRT; 427 AA.
 AC P78807;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-JUN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE FISSION YEAST.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PR745;
 RX MEDLINE; 98162722.
 RA YOSHIOKA S., KATO K., NAKAI K., OKAYAMA H., NOJIMA H.;
 RT "Identification of open reading frames in Schizosaccharomyces pombe
 RT cDNA.";
 RL DNA Res. 4:363-369(1997).
 DR EMBL; D88156; BAA13818.1; -.
 FT SEQUENCE 427 AA; 48056 MW; 47E3B3FE CRC32;

Query Match 2.18; Score 11; DB 3; Length 427;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHHHHHHHHH 12
 |||||
 DB 298 GHHHHHHHHH 308

RESULT 15
 ID 09YGC6 PRELIMINARY; PRT; 467 AA.
 AC 09YGC6;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)

DE METHYL-CPG-BINDING PROTEIN MECP2.
GN MECP2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98282101.
RA JONES P.L., VEENSTRA G.J.C., MADE P.A., VERMAAR D., KASS S.U.,
RA LANDSBERGER N., STROUBOULIS J., WOLFFE A.P.;
RT "Methylated DNA and MeCP2 recruit histone deacetylase to repress
RT transcription.";
RL Nat. Genet. 19:187-191(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA KASS S.U., STROUBOULIS J., WOLFFE A.P.;
RT "Xenopus MeCP2.";
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF106951; AAD03736.1; -.
DR EMBL; AF051768; AAD02651.1; -.
SQ SEQUENCE 467 AA; 51757 MW; A5F353E3 CRC32;

Query Match 2.1%; Score 11; DB 13; Length 467;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 HHHHHHHHNS 13
| | | | | | | | | |
DB 367 HHHHHHHHNS 377

Search completed: August 22, 2000, 17:46:22
Job time: 311 sec


```

89 rleuthrleuglyglytyrlystrhrghlythrthserthservalasnp 106
   |||
210 GCTTACTCTGGCGGTATATAACAGAACCACTAGACATCTGTACT 239
106 hethraspalaalaglyasprometyleuthrthphetrthserglasp 122
   |||
260 TTACAGATGCCGGGTGATCCATGCTACTTAACTTACTTCTCAGCAT 309
123 glyasnasnhslnpetherthrlyvalilleglylyspaserargas 139
   |||
310 GGAATATACCAACCAATCACTACAAAAGTATGGCAAGATCTTAGAGA 359
139 pheaspiliserprolyvalasnnglyluasnlevalglyaspaspy 156
   |||
360 TTTTGTATCTCTCTAAGGTAAACGGTGAGAACCTGTGGGGATGACG 409
156 alvalleualatrhglyserglasphepheyalarserileglyser 172
   |||
410 TCCTCTTGCTACGGGACCCAGGATTTCTTGTCCCTCAATTGGTTCC 459
173 lysglyglylyleualaalaglylystrthrthaspalaalathrvalth 189
   |||
460 AAGCGGTAACTGCGAGGTAATACATGATGCTGAACCGTAAC 509
189 rvalsersnngln.....gluphemetilexylatyrtygluina 203
   |||
510 CGTATCTAACCAAGATCCATCGAACGTCGTATTAGAGCTTACGACACAA 559
203 snproglinhispheilegluasplenglulysvalargvalgluglu 219
   |||
560 ACCACACACATTTTATTGAGGATCTAGAAAAAGTAGGTTGGAACAATT 609
220 thrghlyhsiglysersevalleuglualeuvalleuvallylysas 236
   |||
610 ACTGTCATGCTCTTCAGTTTGAAGAATGGTTCAGTTAGCAAGA 659
236 plyasnilleaspliserilelystrthraspserglu 253
   |||
660 TAAATATATGATATTTCCATTAATATGATCCAGAAAAGATTCGAGG 709
253 alphaealaenargvalillethraspaspilleglualeuvallyly 269
   |||
710 TTTTGCATAGATGATATCTGATATATCCAAATGCTCAGAAAATC 759
270 leualatyrpheleuprogluasprthilleuylslylysthyas 286
   |||
760 CTAGCTATTTTCTACCCGAGATGCATCTTAAAGCGGTCAATTATGA 809
286 paenngluennglylylelysaargvallysglupheuglusers 303
   |||
810 CAACCAACTGCAAAATGCGATCAAGCAGTAAGAGATTCCTGTAACAT 859
303 erfproasntrhglntrpglualeuargalaphemetalaalmethsph 319
   |||
860 CGCGGATATACAAATGCGGATTCGGGCTTCATGCAAGTAATGCAATTC 909
320 serleuthralaaspargilleaspaspillegluvalillevalas 336
   |||
910 TCTTTAACCCCGCATCTGATGATGATATTTTGAAGTATGTTTGA 959
336 psrmetasnhsiglyaspalaargserlyleuarglualeu 353
   |||
960 TTTAAATGAAATCATCATGATGATGATGATGATGATGATGATGATGAT 1009
353 lacluleuthralaglualeuylserileyservalileglnalagluile 369
   |||
1010 CTAGCTTACCGCGCAATTAAGATTTATTCAGTTATTCAGCGCAATTT 1059
370 asnlyshisleserleuylthilleasnillehisaplyseril 386
   |||
1060 AATATACATCTGTCTAGTAGTGACACATTAATATCATGATTAATCAT 1109

```

```

386 easnleumetasplyasnleuylrtyrthrthaspglugluilephei 403
   |||
1110 TAACTCATGATGATAAAATTTATATGTTATACAGATGAGAATTTTAA 1159
403 ysalaaseralaglytyrlylleugluilysmetproglnthrthille 419
   |||
1160 AACCGACCGAGAGTCAAAAATTCGAGAAAATGCTCAAAACCACTT 1209
420 glnvalaspelysergluylslyllevalserillelyaspheleu 436
   |||
1210 CAGGTGATGAGGACGAGAAAAAATAGCTCGATTAAGAGACTTCTTGG 1259
436 ysergluaslyargthrghlyalaleuuglyasnleulysasnsettyrs 453
   |||
1260 AAGTGAGAAATAAAGAACCGGGCGTGGGTATTCGAAAAAATCTACT 1309
453 ertyasnlysaaspaasnnglualeuSerhisphealathrthrysser 469
   |||
1310 CTTATATTAAGATTAATATGATATATCTCATCTTCCACACCTCTCG 1359
470 asplyserargproleuasnaspleuvalsergluylstrthrghlne 486
   |||
1360 GATTAAGTCCAGCGGCTCAACGACTTGTTAGCAAAAAACAACACTGCT 1409
486 useraapilthrserargpheasnseralalleglualaleuasnarap 503
   |||
1410 GTCTGATATTACATCAAGCTTTTATTCAGCTATTGAAACGACGACCT 1459
503 hellegluylstrthraspservalmetglarpheuaspasprthser 519
   |||
1460 TCATTCAGAAATATGATTCAGTATGCAACGTCGTCAATGACAGCTCT 1509
520 glylys 521
   |||
1510 GGTAAA 1515

```

seq_name: gb_pat:A56808

seq_documentation_block: 1462 bp DNA

LOCUS A56808

DEFINITION Sequence 16 from Patent WO9628551.

ACCESSION A56808

VERSION A56808.1 GI:3712821

KEYWORDS

SOURCE

ORGANISM

Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.

REFERENCE 1 (bases 1 to 1462)

AUTHORS Tlthball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,

Bennett and Alice,M.

TITLE VACCINES FOR PLAGUE

JOURNAL Patent: WO 9628551-A 16 19-SEP-1996;

SECOR DEFENCE (GB)

Other publication ZA 9602036 960716

Other publication AU 4951196 961002.

FEATURES

SOURCE

CDS

/organism="Yersinia pestis"

/db_xref="taxon:632"

<8..1450

/note="unnamed protein product"

/codon_start=1

/transl_table=1

/protein_id="CA03421.1"

/db_xref="GI:3712822"

/translation="ADTFASTTATATLVEPARITLTYKEGAPITMDNGNIDELLVG

TLTIGGKRTGTSVNFDTAAGDPMILFTSODGNHOFRTKVGKDSRDFDISPY

NGENLVGDVVATGSDPFEVRSIGSGKLAAGKXDAYTVVSNOSGSEGRIRAYE

QNPQHFIEDLEKVEVQELTGHSVLELVOLVKNIDISIKDPRKDSVFANRYT

TDDELKRTIAYFLPDAILKGGHYNOIOLNGTKRKVEPLESPNOMPLRAFMAY

HFSRTARIDDDILKIVDSNNHGDARSKIRELAETAEKITYSTIQAEINKHLS

SGTINHDKSINLMDKNLYGTTDEIFKASAEYKILKMPOTTIOVDSSEKKIVSIND

FLSSEKRTGALGNLKNYSYNNKDNLSHPATCTSPKSPRLNDVSOKTQQLSDITS
 RFSNLSALNRFIOKXDSVMQRLDDISGR
 BASE COUNT 476 a 286 c 300 g 400 t
 ORIGIN

alignment scores:
 Quality: 2387.50 Length: 482
 Ratio: 5.005 Gaps: 1
 Percent Similarity: 98.963 Percent Identity: 98.340

alignment_block:
 US-08-699-716a-2 x A56808 ..

Align seg 1/1 to: A56808 from: 1 to: 1462

```

43 AsnAlaIAspLeuThrAlaSerThrAlaThrAlaThrLeuValG 59
.....
2 AcCTGGCGAGATTAACTGCAAGCACTGCAACGGCACTGTGTTGA 51
59 UProAlaArgIleThrLeuThrTyrLysGluGlyAlaProIleThrIleM 76
|||||
52 ACCAGCCCGCATCCTTACATATATAGAGAGCGCTCCAAATTACATTA 101
76 eLAspAsnGlyAsnIleAspThrGluLeuValGlyThrLeuThrLeu 92
|||||
102 TCGACAAATGGAACATCGATACAGATTAAGTGTGTCAGGCTACTCTT 151
93 GlyGlyTyrLysThrGlyThrThrSerThrSerValAsnPhetThrAsp 109
|||||
152 GCGCGCTATAAACAGAACCACTACACATCTGTAACTTACATGACATGC 201
109 aAlaGlyAspProMetTyrLeuThrPhetThrSerGlnAspGlyAsnAsn 126
|||||
202 CCGGGGTATCCCATGTACTTAACATTTACTTCTCAGGATGCAAAATACC 251
126 IsGlnPhetThrThrLysValIleGlyLysAspSerArgAspPheAsp 142
|||||
252 ACCAATTCACACAAAGATGATGGCAAGATTCAGAGATTTGATATC 301
143 SerProLysValAsnGlyLysAsnLeuValGlyAspAspValLeuAl 159
|||||
302 TCTCCTAAGGTAAACGTTGAGAACCTTGTGGGGATGACGTCCTTGGC 351
159 aThrGlySerGlnAspPhePheValArgSerIleGlySerLysGlyL 176
|||||
352 TACGGGACCCAGGATTCCTTGTGCTCAATGCTCCTCAAGGGCGTA 401
176 yLeuAlaIleGlyLysTyrThrAspAlaValThrValThrValSerAsn 192
|||||
402 AACTTCACACAGGTAAATACAGCATGCTGTAAACCTTAACCGATTAAC 451
193 Gln.....GluPheMetIleArgAlaTyrGlyGlnAsnProGlnH 206
|||||
452 CAGAGATTCATCGAAGGTGCTATTAGACCTAGCAACAAACCCACACA 501
206 sPheIleGluAspLeuGluLysValArgValGluGlnLeuThrGlyHisG 223
|||||
502 TTTTATTAGGATCTAGAAAAGTTAGGGTGAACAACCTTACGTCACG 551
223 ySerSerValLeuGluGluLeuValGlnLeuValLysAspLysAsnIle 239
|||||
552 GTTCTTCATTTTAGAAGATTGCTCAGTACTCAAAATATAAATAATA 601
240 AspIleSerIleLysTyrAspProArgLysAspSerGluValPheAlaAs 256
|||||
602 GATATTTCATTAATATATATCCAGAAAAGATTCCGAGGTTTTCCTCAA 651
256 nArgValIleThrAspIleGluLeuLeuLysLysIleLeuAlaTyrP 273
|||||
652 TAGAGTAATTACGATGATGCAATGCTCAAGAAAATCCTAGCTTATT 701
273 hLeuProGluAspThrIleLeuLysGlyGlyHisTyrAspAsnGlnLeu 289

```

```

|||||
702 TTCTACCCGAGATCCATCTTAAAGCGGTCATTAAGAACCAACTG 751
290 GlnAsnGlyIleLysArgValLysGluPheLeuGluSerSerProAsnTh 306
|||||
752 CAAAATGGCATCAAGGAGTAAAGAGTTCCTTGATATCGCCGAATAC 801
306 rGlnTrpGluLeuArgAlaPheMetAlaValMetHisPheSerLeuThrA 323
802 ACAATGGGATTCGGGGGCTTCATGCGAGTAATGATTTCTCTTAACG 851
323 lAspArgIleAspAspIleLeuLysValIleValAspSerMetAsn 339
852 CCGATCGTATCGATGATGATTTTGAAGAGTGTGTTGATCATGAT 901
340 HSHISGlyAspAlaArgSerLysLeuArgGluGluLeuAlaGluLeuTh 356
902 CATCATGGGATGCCCGTAGCAAGTTGGTGAGATTAAGTACGACCTTAC 951
356 rAlaGluLeuLysIleTyrSerValIleGlnAlaGluLeuLysHisL 373
952 CCGCAATTAAGATTTATTCAGTTATTCAGCCGAATTAATTAAGCATC 1001
373 eUerSerSerGlyThrIleAsnIleHisAspLysSerIleAsnLeuMet 389
1002 TGTCTAGTAGTGCCACCAATAATATCCATGATTAATCATTAATCATG 1051
390 AspLysAsnLeuTyrGlyTyrThrAspGluGluIlePheLysAlaSerAl 406
1052 GATAAAAATTAATGATTATACAGATGAGAGATTTTAAAGCCAGCGC 1101
406 aGluTyrLysIleLeuGluLysMetProGlnThrThrIleGlnValAsp 423
1102 AGAGTACAAATATCTCGAAGAAATCTCTCAACCACTTCAGTGGATG 1151
423 ySerGluLysLysIleValSerIleLysAspPheLeuGlySerGlnAsn 439
1152 GGAGCGAGAAAATAAGTCTGATTAAGAGACTTCTTGAAAGTAGAT 1201
440 LysArgThrGlyAlaLeuGlyLysAsnLeuLysAsnSerTyrSerTyrAsn 456
1202 AAAAGAACCGGGCGTGGTATATCTGAATAAATCTATCTTAATAATA 1251
456 sAspAsnAsnGluLeuSerHisPheAlaThrThrCysSerAspLysSerA 473
1252 AGATATATATGATTAATCTCACTTTCACCAACCACTGCTGGATTAAGTCA 1301
473 rGProLeuAsnAspLeuValSerGlnLysThrThrGlnLeuSerAspIle 489
1302 GCGCGCTCAGACACTTGGTATGCCAAAAAACACTCAGCTGTCTGATAT 1351
490 ThrSerArgPheAsnSerAlaIleGluAlaLeuAsnArgPheIleGln 506
1352 ACATACAGCTTAATTAATCACTATGACACAGCAACCTGTTTATTCAGAA 1401
506 sTyrAspSerValMetGlnArgLeuLeuAspAspThrSerGlyLys 521
1402 ATATGATTCAGTATGACCAAGCTGTCTAGATGACAGCTGTGTAA 1447

```

seq_name: gb_dal:YEP_LCR

seq_documentation_block:

LOCUS YEP_LCR 2100 bp DNA BCT 26-APR-1993

DEFINITION Yersinia pestis lcrG, lcrV, and lcrH genes, complete cds.

ACCESSION M26405

VERSION M26405.1 GI:155448

KEYWORDS lcrG protein; lcrH protein; lcrV protein; V antigen.

SOURCE 75kb virulence plasmid.

ORGANISM Yersinia pestis

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.

REFERENCE 1 (bases 1 to 2100)

AUTHORS Price,S.B., Leung,K.Y., Barve,S.S. and Straley,S.C.

TITLE Molecular analysis of *lcrGH*, the *V* antigen operon of *Yersinia*
 JOURNAL *J. Bacteriol.* 171, 5646-5653 (1989)
 MEDLINE 9000806
 COMMENT Computer readable copy of sequence [*J. Bacteriol.* (1989) In press]
 kindly submitted by Price, S.
 07-AUG-1989.

FEATURES
 source location/Qualifiers
 1..2100
 /organism="Yersinia pestis"
 /strain="KIM5"
 /db_xref="taxon:632"

-35_signal
 -10_signal
 gene

CDS

147..152
 174..179
 192..479
 /gene="lcrG"
 192..479
 /gene="lcrG"
 /codon_start=1
 /transl_table=11
 /protein_id="AA027640.1"
 /db_xref="GI:155449"
 /translation="MKSSHDEYDKTLKQAEIADSDHRAKILQEMCADIGLPEAV
 KIFAGRSSEEEKPAPRELDDEIKRQREGRPHPDGRPRKPTMMRQIT"
 463..466
 /gene="lcrG"
 463..1461
 /gene="lcrV"
 481..1461
 /gene="lcrV"
 /codon_start=1
 /transl_table=11
 /protein_id="AA027641.1"
 /db_xref="GI:155450"
 /translation="MIRAYEQNPQHIEDLEKRYEQLGHSVLEELVOLVKDKNI
 DISIKYDPKSEVFAANRYITDDIELKKILAYFPEDAILKGHYDQLOMGIRVK
 EFLSSPTOMELRAFMAMVMSRLADRIDDLITVYDSMVRHGDASKEPEELAE
 TRAEKIVYIOAEIKKHLSSTGTTINHDKSTLMKRNLYGTDEITFASAEYKILEK
 MPQTIQVDSSEKKIVSIKDFGSENKRTGALGNLNSYSYKNNKNSHEFATTCDSK
 SNPLDLVQSKTQOLDITSRNSAIEALNREIQRXYDSVMORLDDTSGK"

RBS
 gene
 CDS

1465..1468
 /gene="lcrH"
 1465..1980
 /gene="lcrH"
 1474..1980
 /gene="lcrH"
 /codon_start=1
 /transl_table=11
 /protein_id="AA027642.1"
 /db_xref="GI:155451"
 /translation="MOETTTDQEQOLAMESEFLKGGGTITAMLNISDTEOLYSIAF
 NOYSGKEDAKRVQALCVLDHYDSREFLGAGORAMGQDILHISYSYGAINDIK
 EPRPFHAECLQKGEIAEASGLFLAQELIADTKEFEKELSTRVSSMLEAKIKKEM
 EHECYDNP"

RBS
 gene
 CDS

BASE COUNT 688 a 423 c 461 g 528 t
 ORIGIN

alignment_scores:
 Quality: 1646.00 Length: 326
 Ratio: 5.049 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.693

alignment_block:
 US-08-699-716A-2 x YEPLCR ..

Align seg 1/1 to: YEPLCR from: 1 to: 2100

196 MetlleatrgalatrygluglnasnproglinhisphellegluaspLeuGI 212
 |||||
 481 ATGATTAGAGCTTACGAAACCAACACATTTTATTGATCTAGA 530
 212 utysvalargvalgluInleuthrglyhisglyserValleuLug 229

531 AAAAGTTAGGCTGCAACACTTACTGTGATGTTTTCAGTTTGAAG 580
 229 luleuValgluInleuVallyAspLyAsnIleAspIleSerIleLysTYR 245
 581 AATTGCTGCTAGTCAAGAAATGATAATATGATTTTCCATTAAATAT 630
 246 AspProArgLyAspSerGluValPheAlaAsnArgValIleThrAspAs 262
 631 GATCCAGAAAGATTCGAGAGTTTTCGCAATAGATTAATCTGATGA 680
 262 pIleGluInleuLysIleLeuAlaTYRPhelLeuProGluAspThrI 279
 681 TATCGAATTGCTCAGAAATCTTACTTATTTTCCAGGAGATGCA 730
 279 leuLysGlyGlyHisTYRAspAsnGluInleuGlnAsnGlyIleLysArg 295
 731 TTCTTAAGCGGCTCATATGACACCACTGCAAAATGGCATCAAGCGA 780
 296 ValLysGluPhelLeuGluSerProAsnThrGlnTrpGluLeuArgAl 312
 781 GTAAAGAGTTCCTTGATCATCGCCGAATACACATGGGAATGGCGGG 830
 312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAs 329
 831 GTTCATGGCAGTATGATTCATTTCTTTAACCGCGGATCGATCGATGATG 880
 329 spIleLeuLysValIleValAspSerMetAsnHisGlyAspAlaArg 345
 881 ATATTTTGAAGATGATTTGTTGATTCATGATCATGATGATGATGATG 930
 346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTy 362
 931 AGCAAGTGGCGTGAATAATAGCTGAGCTTACCGCGGAATTAAGATTA 980
 362 rSerValIleGluAlaGluIleAsnLysHisLysSerSerGlyThrI 379
 981 TTGAGTTATTCAGCCGAATTAATAGCATCTGTGTAGAGGCGACCA 1030
 379 leAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTYRgly 395
 1031 TAATATCCATGATTAATCATATATCTCATGATTAATTAATATATGAT 1080
 396 TYTThrAspGluGluIlePheLysAlaSerAlaGluTYRlysIleLeuGI 412
 1081 TATACAGATGAAGAGATTTTAAAGCCAGCGAGATACAAATTCCTCGA 1130
 412 utysMetProGlnThrThrIleGlnValAspGlySerGluLysIleLe 429
 1131 GAAATATGCTCAACCACTTACGATGAGTGAGCGCGGAAAAAATAG 1180
 429 alSerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
 1181 TCTCGATTAAGAGACTTCTTGAGAGTGAAGATTAAGAACCGGGCGCTTG 1230
 446 GlyAsnLeuLysAsnSerTYRSerTYRAsnLysAspAsnGluLeuSe 462
 1231 GGTAAATCTGAAAAACATCACTCTTAAATTAAGATTAATCAATATAC 1280
 462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeu 479
 1281 TCACCTTGGCCACCACTGCTCGGATTAAGTCCAGCGCGCTCAACGACTTGG 1330
 479 alSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
 1331 TTAGCCAAAAAACACTCACTGCTGATATTTACATCAGCTTTAATTTCA 1380
 496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTYRAspSerValMetGI 512
 1381 GCTATTGAAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTATGCA 1430
 512 nArgLeuLeuAspAspThrSerGlyLys 521

1431 ACGTCTGCTAGATGACACGCTGTGTA 1458

seq_name: gb_ba2:YPCD1

seq_documentation_block:

LOCUS YPCD1 70305 bp DNA BCT 22-MAR-2000
 DEFINITION Yersinia pestis plasmid pCD1.
 ACCESSION AL117189
 VERSION AL117189.1 GI:5832423
 KEYWORDS chaparrone; cytotoxic effector; IS100; IS1616; IS1617; lcr; low-calcium response; syc; targeted effector; transposase; type III secretion; Y antitoxin; virulence; Ylp; Yop; ysc.

SOURCE
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.

REFERENCE 1 (bases 1 to 70305)
 AUTHORS Karlyshev, A.V. and Wren, B.W.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 70305)
 AUTHORS Baker, S.G. and Mungall, K.
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 70305)
 AUTHORS James, K.D., Parkhill, J., Barrell, B.G. and Rasthriam, M.A.
 JOURNAL Direct Submission
 TITLE Submitted (03-SEP-1999) Yersinia pestis sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA
 E-mail: barrell@sanger.ac.uk DNA supplied by Dr. Andrey Karlyshev and Prof. Brendan Wren, (3). Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT

COMMENT
 Notes:
 Yersinia pestis sequencing at The Sanger Centre is funded by Beowulf Genomics
 Details of Y. pestis sequencing at the Sanger Centre are available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/Y_pestis/)
 CDS are numbered using the following system eg YPCP1.01c. YP (Y. pestis), PCP1 (plasmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.
 The length in codons is given for each CDS.
 Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database.

The position of possible ribosome binding site sequences are given where these have been used to deduce the correct initiation codon.

CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or att) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

FEATURES

source

1. 70305
 /organism="Yersinia pestis"
 /plasmid="pCD1"
 /strain="CO-92 B1ovar Orientalis"
 /db_xref="taxon:632"

repeat_unit

1. 1956
 /note="IS100 element"

gene

88. 1110
 /gene="YPCD1.01"

CDS

88. 1110
 /gene="YPCD1.01"

/note="YPCD1.01, probable transposase, len: 340 aa; putative insertion sequence IS100, identical to corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946, AF074612) (340 aa), fasta scores: opt: 2328 z-score: 3808.9 E(): 0, 100.0% identity in 340 aa overlap. Similar to many others e.g. TRAO_ECOLI (EMBL:X14793), ISTA, E.coli transposase for insertion sequence element IS21 (390 aa) (33.1% identity in 329 aa overlap). Contains P1am match to entry PF00239 recombinase, Site-specific recombinases,

misc-feature

gene

CDS

112. 195
 /gene="YPCD1.01"
 /note="P1am match to entry PF00239 recombinase, site-specific recombinases, score 25.70, E-value 4.8e-06"
 1110. 1889
 /gene="YPCD1.02"
 1110. 1889
 /gene="YPCD1.02"
 /note="YPCD1.02, probable transposase, len: 259 aa; putative insertion sequence IS100, identical to corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946, AF074612) (259 aa), fasta scores: opt: 1658 z-score: 2693.4 E(): 0, 100.0% identity in 259 aa overlap. Similar to many others e.g. ISTA_ECOLI (EMBL:X14793), ISTA, E.coli transposase for insertion sequence element IS21 (265 aa) (47.4% identity in 249 aa overlap). Contains PS00017 ATP/GMP-binding site motif A (P-loop)"
 /codon_start=1
 /transl_table=11
 /label="YPCD1.02"
 /product="putative ATP-binding protein"
 /protein_id="CAB54879.1"
 /db_xref="GI:5832425"

misc-feature

misc-feature

gene

CDS

1356. 46271
 /note="identical to Y.pestis KIM5 plasmid pCD1 (EMBL:AF053946) from 46489 to 1955, and to Y.pestis KIM5 plasmid pCD1 (EMBL:AF074612) from 59097 to 14563, except where noted"
 complement(1956..2204)
 /gene="YPCD1.03c"
 complement(1956..2204)
 /Partial

/gene="YPCD1.03c"
 /note="YPCD1.03c, probable transposase remnant, len: 83 aa; similar to several e.g. N-terminus of TRAO_ECOLI (EMBL:X14793), ISTA, E.coli transposase for insertion sequence element IS21 (390 aa), similarity is interrupted by the adjacent IS100 element. The remainder of this CDS is in YPCD1.97c. This region is also similar to TR-068707 (EMBL:AF053946, AF074612) Y.pestis KIM5 putative transposase in pCD1 (390 aa). Contains probable helix-turn-helix motif at aa 19-40 (Score 1833, +5.43 SD)"
 /codon_start=1
 /transl_table=11
 /label="YPCD1.03c"
 /protein_id="CAB54880.1"
 /db_xref="GI:5832426"
 /translation="MISREDFYMIKROGAYIIDAIVGCSERTVARYLKPEPP AKTRHKWKLKPFMDYIDIRLAENVNNSVILAIKAM"
 2304. 2948

CDS
 /gene="YPCD1.04"
 2304..2948
 /partial
 /gene="YPCD1.04"
 /note="YPCD1.04, possible transposase remnant, len: 215
 aa; similar to many e.g. TR:Q4612 (EMBL:X78052)
 Enterobacter agglomerans IS 1222 ORF8 (276 aa). Truncated
 at N-terminus"
 /codon_start=1
 /transl_table=11
 /label="YPCD1.04"
 /protein_id="CAB54881.1"
 /db_xref="GI:5832427"
 /translation="HRSGLVHNKRVYRLVHSLGKVRKRRRGLATERPLRPAA
 PNLWSDVDFDALATGRRKICLCVDYRECLTVYAFGISGVQVRIIDSLER
 GPATIRIDQPEFTGCRALDQAFHEHVELIOPGKTQNGFIIESFNGRDECLNE
 HMFSDVSHARKTISEMRDYNCRPHSTLNTOTSEFAAWRKNSDESGDINK"
 2304..2982
 misc_feature
 /note="similar to E.agglomerans IS 1222 (EMBL:X78052) at
 DNA level"
 complement(3014..3406)
 /gene="YPCD1.05c"
 complement(3014..3406)
 /gene="YPCD1.05c"
 /note="YPCD1.05c, syce, yerh, yope targeting protein, len:
 130 aa; identical to corresponding CDS from Y.pestis KIMS
 pCD1 (EMBL:AF053946, AF074612) (130 aa), fasta scores;
 opt: 861 z-score: 1608.5 E(): 0, 100.0% identity in 130 aa
 overlap and to YERA_YERPE (EMBL:M34279) from Y.pestis
 plasmid pYV019. Highly similar to TR:Q56910 (EMBL:Z18539),
 syce, from Yersinia enterocolitica (130 aa) (99.28
 identity in 130 aa overlap) and YERA_YEREN (EMBL:M4278),
 yerh, also from Y.enterocolitica (130 aa) (98.5% identity
 in 130 aa overlap). Similar to TR:Q51448 (EMBL:Z27629),
 ORF1, pseudomonas aeruginosa exoenzyme S ORF1 (116 aa)
 (44.3% identity in 115 aa overlap)"
 /codon_start=1
 /transl_table=11
 /label="syce"
 /product="putative yope chaperone"
 /protein_id="CAB54882.1"
 /db_xref="GI:5832428"
 /translation="WYSREDAITQLFQOLSLSDPTIPYGVKVGACHTHPVQ
 QILMFTPLSDNNDEKELTSLNITFSQILPILSWDEGVGHVLMNKPILNSIDNNS
 LYTQLEMLVQGAERLQTSLSLSPRSFS"
 3600..4259
 /gene="YPCD1.06"
 3600..4259
 /gene="YPCD1.06"
 /note="YPCD1.06, yope, outer membrane virulence protein,
 len: 219 aa; identical to corresponding CDS from Y.pestis
 KIMS pCD1 (EMBL:AF053946, AF074612) (219 aa), fasta
 scores; opt: 1409 z-score: 2140.1 E(): 0, 100.0% identity
 in 219 aa overlap and to YOPB_YERPE (EMBL:M34279), YOPB,
 from Y.pestis plasmid pYV019. Highly similar to YOPB_YERPS
 (EMBL:Y00543), yope, from Yersinia pseudotuberculosis
 plasmid p181 (219 aa) (99.3% identity in 219 aa overlap)
 and to YOPB_YEREN (EMBL:M92066), YOPB, from Yersinia
 enterocolitica plasmid pYV (219 aa) (95.9% identity in 219
 aa overlap)"
 alignment_scores:
 Quality: 1646.00 Length: 326
 Ratio: 5.049 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.693
 alignment_block:
 US-08-699-716a-2 x YPCD1/rev ..
 Align seg 1/1 to reverse of: YPCD1 from: 1 to: 70305
 196 MetLLAAGAlATyTgLuGlnAsnProGlnHISPhelIegLInuSPLeuGI 212
 |||||||
 22915 ATGATTAGAGCCCTACGACAAACCAACATTTATTGAGGATCTAGA 22866

212 ulysValArgValGluGlnLeuThrcLysIserSerValLeuGluG 229
 |||||||
 22865 AAAAGTTAGGTGGAGAACACTTACTGGTACGTCTTCTAGTTTAGAAG 22816
 229 lueValAlnueValLysAspLysAsnIleAspIleSerIleLysTy 245
 |||||||
 22815 AATTGGTCACTAGTCAAGAAATATAATATAGATATTTCCATTAATAT 22766
 246 AspPrArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
 22765 GATCCACGAAAGATTCGGAGGTTTTCGCATATAGATATTAATGATGA 22716
 262 pIleGluLeuLeuLysIleLeuAlaThrPheLeuProGluAspThrI 279
 |||||||
 22715 TATCGAATGCTCAAGAAATCTTATTTCTACCCGAGATGCCA 22666
 279 lLeuLysGlyGlyHisTyTAspAsnGlnLeuGlnAsnGlyIleLysArg 295
 |||||||
 22665 TCTTAAAGCGCGTCATTATGACAACTGCAAAATGGCATCAAGCA 22616
 296 ValLysGluPheLeuGluSerSerProAsnThrGlnTrpGluLeuArgAl 312
 |||||||
 22615 GTRAAAGACTTCTCTCAATCATCGCGAATACCAATGGGAATTCGCGGC 22566
 312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
 22565 GTTCATGCGCAGTAATGCATTTCTTTAACCCGATCGTATGATGATG 22516
 329 sPLeuLeuValIleValAspSerMetAsnHisHisGlyAspAlaArg 345
 |||||||
 22515 AATTTTGAAGTGAATGTTGATCAATGAATCAATCATGATGATGATCCCGT 22466
 346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTy 362
 22465 AGCAATTCGCGAAGAAATTAAGTACGCTTACCGCGAATTAAGATTTA 22416
 362 rSerValIleGlnAlaGluIleAsnLysHisLeuSerSerGlyThrI 379
 |||||||
 22415 TTCAGTTATTAACGCCGAATTAATTAAGCATCTGCTAGTAGGCACCA 22366
 379 lAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyGly 395
 |||||||
 22365 TAAATATCCATGATTAATCATTAATCATGATGATGATTAATATATGTT 22316
 396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyLysIleLeuGI 412
 |||||||
 22315 TATACAGATGAAGACATTTTAAAGCCAGCCAGATACAAATTCCTCGA 22266
 412 ulysMetProGlnThrIleGlnValAspGlySerGluLysIleVal 429
 22265 GAAATGCGCTCAACCCATTCAGGTGATGGAGCGGAAATAATATG 22216
 429 aISerIleLysAspPheLeuGlySerGluAsnLysArgThrLysAlaLeu 445
 |||||||
 22215 TCTCGATTAAGACCTTCTTGAAGGAGAAATAAAGAACCGGCGCTGG 22166
 446 GlyAsnLeuLysAsnSerTySerTyAsnLysAspAsnAsnGluLeuSe 462
 |||||||
 22165 GGTAAATCTGAAAAACATCACTCTTATATAAAGATAATATGAATATATC 22116
 462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
 |||||||
 22115 TCACCTTGCCACACCTGCTCGGATAAGCCAGCGCGCTCAACGACTGG 22066
 479 aISerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
 |||||||
 22065 TTAGCCAAAAAACACTCACTGCTGATATTAACATCACTGTTTATATCA 22016
 496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyTAspSerValMetGI 512
 |||||||
 22015 GCTATTGAAGCACTGACCGCTTCATTCAGAAATATGATTCAGTATGCA 21966

512 naxgleuleaspsthrserglyls 521
 |||||
 21965 ACCTCTGCTAGATGACACGCTGTTAAA 21938
 seq_name: gb_ba2:AF053946
 seq documentation block:
 LOCUS AF053946 70504 bp DNA circular BCT 06-OCT-1998
 DEFINITION Yersinia pestis plasmid pCDL, complete plasmid sequence.
 ACCESSION AF053946
 VERSION AF053946.1 GI:2996222
 KEYWORDS
 SOURCE Yersinia pestis.
 ORGANISM Yersinia pestis
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
 REFERENCE 1 (bases 1 to 70504)
 AUTHORS Hu, P., Elliott, J., McCreedy, P., Skowronski, E., Gaines, J., Kobayashi, A., Brubaker, R. R. and Garcia, E.
 TITLE Structural organization of virulence-associated plasmids of Yersinia pestis
 JOURNAL J. Bacteriol. 180 (19), 5192-5202 (1998)
 MEDLINE 98422474
 REFERENCE 2 (bases 1 to 70504)
 AUTHORS Hu, P., Elliott, J., McCreedy, P., Skowronski, E., Gaines, J., Kobayashi, A., Carrano, A. V., Brubaker, R. and Garcia, E.
 TITLE Direct Submission
 JOURNAL Submitted (16-MAR-1998) Biology and Biotechnology, Lawrence Livermore National Lab, L452, 7000 East Ave., Livermore, CA 94550, USA
 FEATURES
 source Location/Qualifiers
 1. 70504
 /organism="Yersinia pestis"
 /plasmid="pCDL"
 /strain="kim"
 /db_xref="taxon:632"
 87. 1109
 /codon_start=1
 /transl_table=11
 /product="transposase"
 /protein_id="AAC62546.1"
 /db_xref="GI:2996223"
 /translation="MVFETVMEIKILHKGSSRAIAELGISRNTVRYLQANSEP
 KSTVPEPAVASLDEEDYRIQRIADAHYKIPATYIAEIRDOGYRGMTILRAIR
 SYVPOEPAVAREFERPGROMOVGTMRNGSPHFAVAVIGSRMYIETDMR
 YDTLECHRNAREFGVREVLYDMKATVYIQORDVYQGRHFSLSMQFKEMFS
 PRCHPRFRAQTKKVRMYQYTRNSFTILMRLRPMGITVDYETANRGLMHLHVA
 NORKHETIQARPCDRWLEQOSMLALPPEKEDVHLDENLVNFDKHLHPLSYDS
 FCRGVA"
 1109. 1888
 /codon_start=1
 /transl_table=11
 /product="transposase"
 /protein_id="AAC62557.1"
 /db_xref="GI:2996234"
 /translation="MMELOHRLMALAGOLESLISAAPALSOQAVDOESYMDLE
 HILHEKILARHOKAMYRMAAPAKTFEEYDPFAGAPOKOLOSRSSTIERN
 ENIVLPGSGVGTHTLAIMGIEAVAGIKVRTTADLLDLSIQRGGRKTTIQ
 GWAPRLIIDEIGYLPFSQEKLFQVIATKYESAMILTSNLPFGQWDPTFAGDA
 ALTSAMLDRILHSHVQIKGESYRLRKRKAGVIAENPE"
 complement(1939. 2343)
 /codon_start=1
 /transl_table=11
 /product="transposase"
 /protein_id="AAC62568.1"
 /db_xref="GI:2996245"
 /translation="MTEIOASERGGCRINGISRLHICPNTARDPVVEVLOKLAHQ
 YPAGFGLMNLROSGLPNNVRYRYRLKLNRRRGRKRLPRHQPLAIPLM
 NHCWSDPMDALDGRRRRLVVEVILFVNDG"
 complement(2279. 2645)
 /gene="lcrs"
 /complement(2379. 2645)
 /gene="lcrs"
 CDS
 gene
 CDS

/codon_start=1
 /transl_table=11
 /product="low calcium response locus protein S homolog"
 /protein_id="AAC62579.1"
 /db_xref="GI:2996256"
 /translation="MKKARFTETQILRLAKVEGGRRHKDYCRENGSEASYMKSK
 TGGMSDICKMKREERERKRWKMTASLSIDHETLKDVAKKL"
 complement(3193. 3540)
 /gene="yscM"
 complement(3193. 3540)
 /gene="yscM"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein M homolog"
 /protein_id="AAC62586.1"
 /db_xref="GI:2996263"
 /translation="MKINTLOSILNQITQVGHGGQAGRLTETNPITNSHQISAEK
 AFANVLEHVNTALSRHDICLPRVSNLELKQKAGEVITVGTLEQLSDPKLL
 LEAARQOTADG"
 complement(3765. 4430)
 /gene="yscL"
 complement(3765. 4430)
 /gene="yscL"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein L homolog"
 /protein_id="AAC62596.1"
 /db_xref="GI:2996273"
 /translation="MSQTCQGTGYAMPTFOVITNSLSLACGLILRAEDYQSITTE
 ELISAQKQDAKILADAGVEYEQKQLQWQMGMDARTLQTLHETOLQOQFYRHV
 EQMSEVVLAVRKILNDQVAMTQVRALEALVSNQGVAVRVNAGALREHY
 AKVHKDFEISYLEVTADARLDQGGCILETVGIIDASIDQIALSNRISTLTGQMK
 VTE"
 complement(4376. 5005)
 /gene="yscK"
 complement(4376. 5005)
 /gene="yscK"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein K homolog"
 /protein_id="AAC62605.1"
 /db_xref="GI:2996282"
 /translation="MKNYITSPFLRCPAAVYILHQLPSIMRSILPYLPQWRDASHL
 NNAIDESTLQIDEEPHGLGALPQPOSQELICRLGLVHGAIRCVLASPIQ
 LLTYNQSLQITVOHEHLLGPNPTHQRLPRIEBRTYQSGIAFWLAAMEPOQ
 AMCKNLSTRLPATESEPMVAESQRLAQTICHLVQVPTCSHLEK"
 complement(5005. 5739)
 /gene="yscJ"
 complement(5005. 5739)
 /gene="yscJ"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein J"
 /protein_id="AAC62607.1"
 /db_xref="GI:2996284"
 /translation="MKVTSITLILFLTGCVVDYLTGISQKGNMALLAQEGL
 SADKPDQSGIKILVSESDVAQADIDIKRKGYSSTIQDVPKRGILSSPDEL
 ARLNARAKQESITLSEIDGVLAARVLEEDNNRKRKVAASAVFIHADIQF
 DTYLPIQIVLNNSIEGLAYDRISVILVPSVYROSSHLPRNTSILTSIOVSESGHL
 IGLSILLPLPVTLAQYFWLQRRK"
 complement(5746. 6093)
 /gene="yscI"
 complement(5746. 6093)
 /gene="yscI"
 /codon_start=1
 /transl_table=11
 /product="Yop proteins translocation protein I homolog"
 /protein_id="AAC62608.1"
 /db_xref="GI:2996285"
 /translation="WPNTEIAQADEVIITTEIGAPAEPTDQIMRDAENSEDTQGL
 GSHLKEVSDIOKSEFTYKSDLHTRKLAVSVDNPDMLMOMSLIRITQIEBLIAKTAG
 RMSQNVETLSRGG"
 complement(6094. 6591)
 gene

```

/ gene="ysch"
/ complement(6094. .6591)
/ gene="ysch"
/ codon_start=1
/ transl_table=11
/ product="Yop proteins translocation protein H homolog"
/ protein_id="AAC62547.1"
/ db_xref="GI:2996224"
/ translation="MTVTLLKRGSTLSMSSQAVSTLQPVASELKTQLENKKSSEAE
KTRVELMOQYYASNPDPHAEVLAEPVREALLARGQHGQSVPAIDLPRLSVLQO
FDSFGKRWEMAILIQVLEGIKPNESQVGLPYSELINKEMILPNSIYDSILHNSHO
IDMDT"
/ complement(6588. .6935)
/ gene="yscg"
/ complement(6588. .6935)
/ gene="yscg"
/ codon_start=1
/ transl_table=11
/ product="Yop proteins translocation protein G homolog"
/ protein_id="AAC62548.1"
/ db_xref="GI:2996225"
/ translation="MKYKLVLLAEIALIGTGNHYHEANCIAMWHLKGEERAVOLI
RLSSLMNRGDYASALQGNKLAIPDLEPMIALCEYRLGLGSALESRLNRLARSQDPRI
QRFVNGMRBQLKT"
/ complement(6937. .7200)
/ gene="yscf"
/ complement(6937. .7200)
/ gene="yscf"
/ codon_start=1
/ transl_table=11
/ product="Yop proteins translocation protein F homolog"
/ protein_id="AAC62549.1"
/ db_xref="GI:2996226"
/ translation="MSNFSGETGTDIADLDAVAQTLKKRADANKAVNDSIALKDK
PNDPALADLQHSINKMSVYININSTIVSRMSKMDLMOGIIQKFP"
/ complement(7201. .7401)
/ gene="ysce"
/ complement(7201. .7401)
/ gene="ysce"
/ codon_start=1
/ transl_table=11
/ product="Yop proteins translocation protein E homolog"
/ protein_id="AAC62550.1"
/ db_xref="GI:2996227"
/ translation="KTQLEQHLNVTWVSITNMLEMALTKLKKDMRGSDAKQYQVW
QRESKALSAIAIHYVAGDLK"
/ complement(7398. .8657)
/ gene="yscd"
/ complement(7398. .8657)
/ gene="yscd"
/ codon_start=1
/ transl_table=11
/ product="Yop proteins translocation protein D homolog"
/ protein_id="AAC62551.1"
/ db_xref="GI:2996228"
/ translation="MSWVCRFYQGRHGVVELPHRCVYFGSDPLQSDIVLSSEIAP
VHYLVWDEGRITDLSAEPLDEGLPVPGLTLKAGSCLEVGFLMTVAVGQPIPE
TLDQVPRKEPDLRPSRLGIGVLSLLLTFLGLHGMREYNDDQGVDEBEY
RRLATAAKDVVLTSPKKEGEPMLTGYIODNHARLSIONFLESIGIFPRLERAME
ELRQGAFFLQRLGRIYGIYSLAPQAGMLOLNEVEEELQKIDSLDAEYGLGIV
ESKVRIRAGNQRKRLDALLBQGLDSDFTYVVKELLELNGOVNDEKLNFSNQQLFR

```

alignment_scores:
 Quality: 1646.00 Length: 326
 Ratio: 5.049 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.693

alignment block:

US-08-699-716a-2 x AF053946 ..

Align seg 1/1 to: AF053946 from: 1 to: 70504

196 MetleatgataTyrGluGlnasnrpGlnHspheilegluaspleucl 212

```

|||||
25311 ATGATTGAGCCTACGACCAACACCATTTTATGTGGATCTAGA 25360
212 ULyVaLaRgYaIgluGlnLeuThrGlyHisGlySerSerValLeuGluG 229
25361 AAAAGTTAGGTTGGACACTTACTGTCATGCTTCTTCAAGTTTAAAG 25410
229 lUleuValGlnLeuValIysAspLysAsnIleAspIleSerIleLysTyr 245
25411 AATTGCTCACTGATGCAAAAGATATAATATGATATTCATTAATAATAT 25460
246 AspProaRdLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
25461 GATCCAGCAAAAGATTCGAGAGTTTTCACATAGAGTAAATTAATCTGATGA 25510
262 pIleGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThri 279
25511 TATGAATGCTCAAGAAATCCTAGCTTATTTCTACCGAGAGATCCCA 25560
279 lUleuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
25561 TTCTTAAAGCGCGTCATTAATGACCAACCACTGCAAAATGGCATCAACGGA 25610
296 ValIysGluPheLeuGlnSerSerProAsnThrLntPrpGluLeuArgAl 312
25611 GTAAAGAGTTCCTTGATCATCGCCGAATACACAAATGGGAATGGCGGC 25660
312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAspA 329
25661 GTTCATGCGCATATGCAATTCCTCTTAACCGCGCATGATCATGATATG 25710
329 sPleLeuLysValIleValAspSerMetAsnHisHisGlyAspAlaArg 345
25711 ATATTGGAAGTATGATTCATCAATGATCAATGATGATGATGCCGT 25760
346 SerIysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTy 362
25761 AGCAAGTTCGCTGAAGATTTTAAAGCCAGCAGATTAAGAAATTCCTGA 25810
362 rSerValIleGlnAlaGluIleAsnLysHisLeuSerSerSerGlyThri 379
25811 TTCAATGATTAACCCGAATTAATATAGCATCTCTAGTAGTGCCACA 25860
379 lEaSnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
25861 TAAATATCCATGATTAATCCATTAATCTCATGATTAATAATTTATAGGT 25910
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluLntYrIysIleLeuG 412
25911 TATACAGATGAAGAGATTTTAAAGCCAGCAGATTAACAAATTCCTGA 25960
412 uLyMetProGlnThrThriIleGlnValAspGlySerGluLysLysIleV 429
25961 GAATATGCTTAACCAACCATCTAGTGATGGAGCGGAGAAAAAATATAG 26010
429 aISerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
26011 TCTGATAAAGAGACTTCTTGGAGTAGAATAAAGAACCGGGCGCTTG 26060
446 GlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe 462
26061 GGTATCTGAAAAACATCACTCTTATATAATGAATATATGATATATC 26110
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
26111 TCATTTGCCACACCTGCTCGATAGTACAGGCCCTTAACGACTTGG 26160
479 aISerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
26161 TTACCAAAAAACAATCAGCTGCTGATATATCATCATCAGCTTTAATCA 26210
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG 512
|||||

```

26211 CCTATTGAAGCCTGACCGTTTCATTCAGAAATGATCAGTATGCA 26260

512 nARgleuenaaspRhrserGlylys 521

26261 ACCTCTGCTAGATGACAGCTGTGTAA 26288

seq_name: gb_ba2:AF074612

seq_documentation_block:

LOCUS AF074612 70559 bp DNA circular BCF 07-Apr-2000

DEFINITION Yersinia pestis plasmid pCD1, complete plasmid sequence.

ACCESSION AF074612 M25810

VERSION AF074612.1 GI:3822037

KEYWORDS

SOURCE

ORGANISM

Yersinia pestis.

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Yersinia.

1 (bases 4318 to 44840)

Leung,K.Y. and Straley,S.C.

The yopM gene of Yersinia pestis encodes a released protein having

homology with the human platelet surface protein GPIIb/alpha

J. Bacteriol. 171 (9), 4623-4632 (1989)

89359090

2 (bases 1 to 70559)

Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and

Blattner,F.R.

DNA sequencing and analysis of the low-Ca²⁺-response plasmid pCD1

of Yersinia pestis KIM5

Infect. Immun. 66 (10), 4611-4623 (1998)

98427122

3 (bases 4318 to 44840)

Straley,S.C.

Direct Submission

Submitted (26-APR-1993) Microbiology and Immunology, University of

Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA

4 (bases 1 to 70559)

Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and

Blattner,F.R.

Direct Submission

Submitted (25-JUN-1998) Microbiology and Immunology, University of

Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA

On Apr 7, 2000 this sequence version replaced gi:155539.

Location/Qualifiers

1..70559

/organism="Yersinia pestis"

/plasmid="pCD1"

/strain="KIM5"

/db_xref="taxon:632"

57..368

/gene="Y0001"

57..368

/note="Y0001: 43 pct identical (0 gaps) to 100 residues of

an approx. 200 aa protein GENEPT: gi153126,orf_0198

Escherichia coli"

/codon_start=1

/transl_table=11

/product="unknown"

/protein_id="AAC69758.1"

/db_xref="GI:3822038"

/translation="MHQSRGAGSRTSLRMROSGYVWVRLARRLMRCGLASROP

GKPRYRGREVSASPDLKRFKRPSEPNRWISYIYKNGWCYLAIVIDLYSRH

W"

665..1033

/gene="nuc"

/note="Y0002"

665..1033

/gene="nuc"

/codon_start=1

/transl_table=11

/product="endonuclease"

/protein_id="AAC69759.1"

gene

CDS

misc_RNA

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

/db_xref="GI:3822039"

/translation="MDTKLQHTPIGTWVDRPVNTKSGKRLRCPDFVHYRMDL

VNAGIPVPTVNSFKALHDKVITVDGKNTOMGSEFNFOAIVQNSSEVLLIMGDFVQV

AYLOVWQSRNKKGTWRRSSY"

1171..1425

/gene="repB"

/note="Y0003"

1171..1425

/gene="repB"

/codon_start=1

/transl_table=11

/product="replication protein"

/protein_id="AAC69760.1"

/db_xref="GI:3822040"

/translation="MSQIENAVTSSSKRAYRGNPLTGAQKQMSYRKKETHKAINV

FIQNDKNELQCEGSGTQETEMRWLQREKARTNA"

complement(1560..1643)

/note="antisense RNA"

/product="copA"

1667..1741

/gene="tap"

/note="Y0004"

1667..1741

/gene="tap"

/codon_start=1

/transl_table=11

/product="repa translation protein"

/protein_id="AAC69761.1"

/db_xref="GI:3822041"

/translation="MRKVQYLLRLLLPCNISGRCD"

1734..2600

/gene="repa"

/note="Y0005"

1734..2600

/gene="repa"

/codon_start=1

/transl_table=11

/product="replication protein"

/protein_id="AAC69762.1"

/db_xref="GI:3822042"

/translation="MTNHQALFTHTHYRQVKNPPEPTPEGGKTLPECKRLAKAEGF

TSRFDLSMVAFAFSLSLRHPPLRRRAIDALOGMCFHDPPLNRIQRTNTAI

EGCLATESGNSITRATRALFSLGLIYQVYDQICNIPITDIFPALFSA

LQSDVAFAAARSRVEMQOERKORLREXDELIAAMRFRFRPSYQTERAAH

GLKRAARDVDRTDRDIAIVNQITRIAEGRFGNIDAVREARAVKERMMSR

NNYTRLATGAT"

1741..2935

/note="orfR"

complement(3427..3645)

/gene="Y0006"

complement(3427..3645)

/gene="Y0006"

/note="orf5 (f72): 42 pct identical (0 gaps) to 33

residues of an approx. 216 aa protein GENEPT: gi12055297,

Imp2 Xenopus laevis"

/codon_start=1

/transl_table=11

/product="unknown"

/protein_id="AAC69763.1"

/db_xref="GI:3822043"

/translation="MRSPLAYGGSYTVVVTQIHLNDIKSAVKHKRLVHVLGQASS

AOVFVMPHKGMAKPADIVNVDNMR"

4758..5186

/gene="Y0008"

4758..5186

/note="orf7 (ol142): 31 pct identical (1 gap) to 48

residues of an approx. 104 aa protein GENEPT: gi12149940,

orf1 p. syringae"

/codon_start=1

/transl_table=11

/product="unknown"

/protein_id="AAC69764.1"

```

gene
  /db_xref="GI:3822044"
  /translation="MINTFTELLPKIASHFGLDKLSODEXGLCELLINDRVIMLRA
  DEILNRLILGLFSGPEARSASQLEFYISIALNKDGPFCFAMSEELGLIAFKHL
  SLDELINVENSVSEIANFYDLVSLVSLPAETAPSYISIG"
  5204..7402
  /gene="ypka"
  /note="10009"
  5204..7402
  /gene="ypka"
  /codon_start=1
  /transl_table=11
  /product="secreted protein kinase"
  /protein_id="AAC69765.1"
  /db_xref="GI:3822045"

gene
  /db_xref="GI:3822046"
  /translation="MKSVKIMGTMPSSISLAKAHERISQHMNPVGLINIGKRRRII
  DNOVLRNPHSGFSLFRGVGKIFSGKFNFSIANLDTLAAQKTSQELRSIPN
  ALSNLFKAPOTELPLGKWEPLSGAPLEGRVAVETKFPAGESHSIIIEKDKORL
  VAKIRSTAEHGLFALEAYKHIVYTAGKHPIANVHGMAYVPGNREPEALINDENY
  GMRCDIARTLADSKROCKINSEATWGITKTAHRLDVTNHLAKAGVHNDIKRNV
  VEDRASGEPPVYIDGLHSRSGEPKGFTESEKAPLGVGNLASEKSDVFLVSTLLH
  CIEGKNEPEIKPNGLRFTISEPAHVDENGYPHREGIAGVETARYRTPTLIGVS
  ADSRDSNEARLHEFLSDGITDEESAKQILKDTLGEKSPISDVRYRTPKLIELSD
  LLRHLSSAATKOLDMGVLSPLDTMLVALDKAREGGVDKOLKSFNSILKTYRVI
  EDYVKGREGDKNSTEYSPYHSNFMLSIYEPSPQRIQKHLDOHSDSDIGSLYRAH
  KHLLELVLYLSQGGVSEETIGFLNRLTEAKITLSQQLNTLQOQOSESAAQDLST
  LNRSGSWADVAHQSLQRFSDTOPVAKFTGQYTAIRHOMMAAHAATLLOVSEETD
  MKNFTVDSIPILQLGRSSLMDEHLVEOREKIRELTIAERLNRLEBEM"
  7798..8664
  /gene="yopJ"
  /note="10010"
  7798..8664
  /gene="yopJ"
  /codon_start=1
  /transl_table=11
  /product="targeted effector protein"
  /protein_id="AAC69766.1"
  /db_xref="GI:3822046"

gene
  /db_xref="GI:3822047"
  /translation="MIGPISQINISGSLSEKETSLSINEEIKNITTOLEFIDISDSM
  FHKNSRDVEWPAVLVQANNKYPEMNINLVTSLDLSIEIKNVIENGVSRRILIN
  MEGEGIHFSVIDYKHINGTSLILFEPANFNMGPMALAIRKTAIERYQCPDCHFS
  VEMDIORSSSEGLFSLAKLKIYERDLSLKHIDNIGLIDENPLHDKLDPYL
  PTFEYKHOGSKRRLNEVINTNGVGYVNNKNEIIVNRPNNKSIYDGLKELSVYKH
  KRIAEKTKLKY"
  9838..10064
  /gene="yopI2"
  9838..10064
  /gene="yopI2"
  /note="068: 45 pct identical (0 gaps) to 68 residues of an
  approx. 560 aa protein GENPEPT: g12162435, hypothetical"
  /codon_start=1
  /transl_table=11
  /product="unknown"
  /protein_id="AAC69767.1"
  /db_xref="GI:3822047"

gene
  /db_xref="GI:3822048"
  /translation="NAVGRKNLFPAGSLRAGORASISLLETAKLNHDPYWLKRDY
  LTRLETPNSQNALPLPAENRFS"
  10347..11753
  /gene="yopH"
  /note="10013"
  10347..11753
  /gene="yopH"
  /codon_start=1
  /transl_table=11

```

alignment_scores:
 Quality: 1646.00 Length: 326
 Ratio: 5.049 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.693

alignment_block:
 US-08-699-716a-2 x AF074612 ..

Align seg 1/1 to: AF074612 from: 1 to: 70559

```

196 MetIleArgAlaTyrGluGlnAsnProGlnHisPheIleGluAspLeuG1 212
37919 ARGATTAGAGCCTACGAAACAAACCAACCAACATTTATTATGAGCATCTGA 37968
212 uysValArgValGluGlnLeuThrGlyHisGlySerSerValLeuGluG 229
37969 AAAAGTTAGGGTGAACAACCTACTGCTCATGCTCTTCAGTTTAGAG 38018
229 IuleuValGlnLeuValLysAspLysAsnIleAspIleSerIleLysTyr 245
38019 ANTGGTTAGTAGTACCAAGATATAAATATGATATTTCCATTAAAT 38068
246 AspProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
38069 GATCCACGAAAGATATCGAGAGTTTTCACATATAGATATATGATGATGA 38118
262 pIleGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThrI 279
38119 TATCGAATTGCTCAAGAAATCTAGCTTATTTCTACCCGAGATGCA 38168
279 IeLeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
38169 TTCTTAAAGCGGCTCTTTTACACCAACCTCAAAATGGCTCAAGCA 38218
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTyrGluLeuArgAl 312
38219 GTAAAGAGTTCCTGATATCATCGCGAATATACATAGGATGCGGGC 38268
312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspArg 329
38269 GTTCATGCGCATGATGATCTCTTAAACCGCGCATCATGATGATGAG 38318
329 spIleLeuLysValIleValAspSerMetAsnHisHisGlyAspAlaArg 345
38319 ATATTTTGAAGAGATGTTGATTCATATCATATCATATGATGATGCCCT 38368
346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTy 362
38419 TTCAGTTATTCAGCCGGAATTAATAGCATCTGTCTAGTATGCGACCA 38468
362 rSerValIleGlnAlaGluIleAsnLysHisLysSerSerSerGlyThrI 379
38469 ACAGAACTTCGCGAAGATTAATAGCTGATCGCCGCAATTAAGATTGA 38418
379 IeAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
38469 TAAATTCATGATTAATCATTAATCATATGATGATTAATTAATGAT 38518
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuG1 412
38519 TATACAGATGAAGAGATTTTAAAGCCAGCGAGATACAAATCTCGA 38568
412 uysMetProGlnIleThrIleGlnValAspGlySerGlyLysLysIleV 429
38569 GAAAGGCTCAACACCATTCAGGTGATGGAGCGGAAAAAATAG 38618
429 aLserIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
38619 TCCGATTAAGGACTTCTTGGAGAGAGATTAAGATTAAGATTAAGATTA 38668
446 GluAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe 462
38669 GGTAAATCTGAAAAACTCATCTTATATTAAGATTAATAGATTAATC 38718
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
38719 TCACITTCGCACACCTGCTCGATTAAGTCCAGCGCGCTCAACGACTTGG 38768
479 aLserGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
38769 TTAGCCAAAAAACACTAGCTGCTGATTAATTAATCAACGTTTAAATCA 38818

```

260 RASDASRIEGLEULEULEULSVSTIETLEUA)ATYRPHELLEUPROG)UA 277

Yersinia pestis.

[illegible][illegible]


```

/transl_table=11
/protein_id="CA02894.1"
/db_xref="GI:2300615"
/translation="GIPGIRAYEONPOHFIEDLEKVEEOLTGHSVLEIYOLVND
KNIDISIKYPRDRSEYFANRVTDDLELKILATILYLPEDAILKSGHYNOLONGIK
RVEFELESSPNTOMELRAFAYVHFSITADRIDDDILKYVDSNHHGDRSKIREL
AELTAEIKITSVIOAEINKHLSSSGTINIDKSLNMDKNLYGTDEIFKASAEYKI
LEKMPOTTIOVDSEKIKVISIKDFLSENNRTALGNLKNYSYNDKNDELSPHATTC
SDSRPINDIVSOIKTQLOSDITSRFSNIAELNRFIOKYSVMQRLDDTSRK"
BASE COUNT      343 a      184 c      205 g      282 t
ORIGIN

```

```

alignment_scores:
  Quality: 1641.00      Length: 325
  Ratio: 5.049          Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 99.692

```

alignment_block:

US-08-699-716A-2 x A46413

Align seg 1/1 to: A46413 from: 1 to: 1014

```

197 ILEATGALATYTGUGLNASNPROGHNISPHIELGLUASPLEUGLULY 213
113 TGGTCAGTACTCAAGATAAATAATAGATATTCCTTAATAATGAT 162
247 PROATGLYASPSERGLUVALPHEALASNAAGVALIETHRASPA 263
163 CCCAGAAAAGATTCGAGCGGTTTCCCAATAGCATGATGATAT 212
263 EGLUENLEULYSILILEUALATYRPHLEUPROGHNISPHIEL 280
213 CGAATTCGCAAGAAATCCCTAGCTATTTTCCACCGAGAGCCATTC 262
280 EULYSGLYLNHSTYRASPSANGLNUGLNASNGLYILEYASPA 296
263 TTRAAAGCGGTCATTTGACACCACTCGAAATGCGATCAAGCAG 312
297 LYSGLUPHELEUGLUSERSEPROASNTHRGINTRPGLULEA 313
313 AAGAGATTCCTTGAAATCATCGCGAATACCAATGGGAATGCGG 362
313 EHEIHAVALMETHISPHESERLEUTHRALASPAAGLIEASPA 330
363 CATGGAGAAATGCAATTCCTTTAACCCCGATCCTATCGAGAGA 412
330 LLEULYSVALILEVALASPSERMETASNHISGLYASPAIA 346
413 TTTGAAATGATGTGTTGATTCATATCATCATGCTGATGCCCT 462
347 LYSLEUARGLUGLULEUALAGLULEUTHRALAGLULEULYS 363
463 AAGTTCGCGAAGAAATTTAGCTGAGCTTACCGCGAATTAAGA 512
363 TYALLIEGLINALAGLULIASNLHSTYRASPSERSEGLYTH 380
513 AGTATATCAAGCGGAATTAATAGCATCTGTCTAGTGGCACA 562
380 SNLIEHISAPLYSSERIEASNLMEETASPLYSANLEUTYR 396
563 ATRATCATATTAATCATATTAATCATGATTAATTAATGAT 612
397 THRASPLUGLULIEPHELYSALASERALAGLUTYR 413

```

```

613 ACAGATGAAGAGATTTTAAACCGACGCGAGATACAAATTCGAGAA 662
413 SKEETPROGHNTHRILIEGLINVAISPGIYSEGLUYSILYVALS 430
663 AATGCTCTCAACACCATTCAGTGGATGGAGCGAGAAAAAATAGCT 712
430 ERIELYSAPHELEUGLYSERGIUANLYSARGTHGLYALAEUGLY 446
713 CGATTAAGAGACTTCTTGAAAGTGAATTAAGAACCGGGCGGTGG 762
447 ASNLEULYSASERTY:SERTYRASNLYSAPASANSGLULEUS 463
763 AATCGAAAAACTCATCTATATATTAAGATAATATGATATCTCA 812
463 SPHEALATHRTHYCSERASPLYSERARPROLEUSNAPLEUVALS 480
813 CTTGCCACACCTCTCGGATAGCTCCAGCGCTCAACGCHTGTG 862
480 ERGLNLYSPTHRGINLEUSERASPILETHRSERARPHENASER 496
863 GCCAAAAAACACTCAGCGTGTGATATTAATCACTGTTTATACCT 912
497 IIEGLUALALEUASNAIRGPHIELGLNLYSTYRASPSERVAL 513
913 ATTGAAGCACTGACCGTTTCATTCAGAAATATGATCAGATCA 962
513 GLEULYSAPSPHTRSERGLYLYS 521
963 TCTGCTAGATGACGCTGTGTTAA 987

```

seq_name: gb_pat:A56795

seq_documentation_block:

LOCUS A56795 1014 bp DNA PAT 03-MAR-1998

DEFINITION Sequence 3 from Patent WO9628551.

ACCESSION A56795

VERSION A56795.1 GI:3712810

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

CDS

BASE COUNT

ORIGIN

```

alignment_scores:
  Quality: 1641.00      Length: 325
  Ratio: 5.049          Gaps: 0

```

Percent Similarity: 100.000 Percent Identity: 99.692

alignment_block:

US-08-699-716a-2 x AS6795

Align seg 1/1 to: AS6795 from: 1 to: 1014

```

197 ILAARGALATYRGLUGINASPARGINHPHELEGLUASPLEUGLU 213
13 ATTCGAGCCACGAAACCAACCAACATTTATTAGATCTAGAAAA 62
213 SVALARGVALGLUGINLEUTHRGHLSGLYSERSEVALLEUGLU 230
63 AGTTAGGCTGGAACACTTACTGCTACTGCTCTCTCTCTCTCTCT 112
230 EUVALGLINLEUVALYASPLYSANILEASPILESERILEYSTYR 246
113 TGGTTCAGTTAGTCAAGATAAATAATAGTATTTCCATTAAATG 162
247 PROARGLYSASPERGLUVALPHEALASNARYVALLETHRASPA 263
163 CCCGAGAAATTCGAGGTTTCCCAATAGACTAATCTGATGATAT 212
263 EGLULEULEULYSILEUVALATYRPHLEUPROGLUASPHRIEL 280
213 CGAATTCCTACAGAAATTCCTAGCTATTTCTACCCGAGATGCC 262
280 EULYSGLYGLYHISTYRASPASNGINLEUGINASNGIILEYAR 296
263 TTAAGCGCGGTCAATTATGACACCACTGCATAATGCGATCAAC 312
297 LYSGLUHEUGLUSERSERPROASNTGINTREPLUENALGALAP 313
313 AAGAGTCTTGATATCGCCCAATACCAATGGGAATTCGCGGCTT 362
313 EMETALVALMETHISPHESERLEUTHRALASPARGLEASPAR 330
363 CAHGGCGATATGATCTCTTTACCCGCGATCGATCATATGATA 412
330 LEULYSVALILEVALSPSERMETASNHSISGLYASPALAARG 346
413 TTTTGAAGATGATGTGATTCATGATCATGATGATGATGATG 462
347 LYSLEUARGLUGLULEUVALGLULEUTHRALAGLUENULYSI 363
463 AAGTGGGTGAAGATGATGATGATGATGATGATGATGATGAT 512
363 VALILEGLALAGLULEUVALYASPLYSANILEASPILESER 380
513 AGTTATTCAGAACCCGAAATTTAATAGCATCTGCTAGTAGTG 562
380 SNILEHASPLYSERILEASNULEUVALYASPLYSANILEYGL 396
563 ATATCCATGATTAATTCATTAATCTCATGATAAATAATTAAT 612
397 THRASPLUGLULEUHELYSALASERALAGLUYRYSILEUGLU 413
613 ACAATATGAAGATTTTAAAGCCAGCGAGATGACAAATTCG 662
413 SMEPRGGLTHTRHLEGINVALASPLYSERGLULYSILEVAL 430
663 ATATGCTCAACCCACATTCAGTGGATGGAGGAGAAAAAATAG 712
430 ERILEYASPLHEUGLISERGLUASNUYASARGTHGLYALAGLU 446
713 CGATAAAGACTTCTTGGAAGTGAATAAAGAACCGGCGCTTGG 762
447 ASNLEULYSANSERTYRSETYRASNULYASPARASNGIUL 463
763 AACTGTGAATAACTACTCTTATATTAAGATATATGATATCTCA 812
463 SPHEALATHTRCHYSERASPLYSERARGPROLEUASNSP 480

```

```

813 CTTTGCACCACTGCTGCGATAGTCCAGCCGCTCAGCATTTGTTA 862
480 ERGLNYSRTHRGHLEUASERASPILETHSERARGPHEASERALA 496
863 GCCAAAAACACTGCTGATATTTACATCAGCTTTATTCAGCT 912
497 ILGLUALALEUASNARGPHELEGINLYSTYRASPSEVALMET 513
913 ATTGAGCACTGAACCGTTTCATTCAGAAATATGATGATGACAG 962
513 GLEUASPARSPHRSERGLYLYS 521
963 TCTGCTGATGACACGCTCTGTTAAA 987

```

seq_name: gb_baz:AF167309

seq_documentation_block: 975 bp DNA

DEFINITION Yersinia pestis strain Pestoides F v antigen (icrv) gene, complete

ACCESSION AF167309.1 GI:7578512

VERSION AF167309.1 GI:7578512

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

alignment_scores:

Quality: 1631.00

Ratio: 5.050

Percent Similarity: 100.000

Percent Identity: 99.690

alignment_block:

US-08-699-716a-2 x AF167309

Align seg 1/1 to: AF167309 from: 1 to: 975

196 METLEARGALATYRGLUGINASPARGINHPHELEGLUASPLEUGLU 212

1 ATGATTAGAGCTACGAAACAAACCCACAACTTTATTATGAGAGCTCAGA 50
212 ulysValArgValGluGlnLeuThrGlyHisGlySerSerValLeuGlu 229
51 AAAAGTTAGGGGTGGCAACAACTACCTGGTCATGGTCTTCCAGTTTGAAG 100
229 IuleValGlnLeuValLysAspLysAsnIleAspIleSerIleLysIle 245
101 AATTGGTTCACTTACTCAGAAAGTAAATAATATGATATTTCCATTAAATAT 150
246 AspProArgLysAspSerGluValPheIleAsnArgValIleThrAspAs 262
151 GATCCACAAAAAGATTCGAGAGTTTGGCAATAGGATATTTCTATGTA 200
262 PileGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThr 279
201 TATCGAATTTGGTCCAGAAATCTCGAGCTATTTTTCAACCCGAGAGATCCA 250
279 IeLeuLysGlyIleHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
251 TTTCTTAAAGCGCGTATTATGCAACCACTGCATAATGGCAATCAAGCGA 300
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTrrGluLeuArgAl 312
301 GTAAAGAGTTCTCTTGATCATCGCCGATACACAAATGGGAATTCGGGGC 350
312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAspA 329
351 GTTCATGGCAGTAATGCAATTTCTCTTAAACCGCGCATCGATGATGATG 400
329 spIleLeuLysValIleValAspSerMetAsnHisIleGlyAspAlaArg 345
401 ATATTTTAAAGTGATGTGTGATTCAATGAAATCATCATGATGATGCCCG 450
346 SerLysIleuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyr 362
451 AGCAAGTTCGGGAGAGAAATTACCTGAGCTTACCGCGCAATTAAAGATTTA 500
362 rSerValIleGlnAlaGluIleAsnLysHisLysSerSerSerGlyThrT 379
501 TTTCAGTATTCAAGCGCAATTAATATAGCATCTGTCTAGATGGGCACCA 550
379 IeAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
551 TAAATATCCAGATGATTAATCCATTATCTCATGGATTAAAAATTATATG 600
396 TyrThrAspGluGlnIlePheLysAlaSerAlaGluTyrLysIleLeuGlu 412
601 TATACAGATGAAGAAATTTTAAAGCCACGCGAGATACAAAATTTCTCGA 650
412 ulysMetProGlnThrThrIleGlnValAspGlySerGluLysIleVal 429
651 GAAATGCTCCAAACCACCATTCAGGTGATGGGACGCGAAGAAAAATAG 700
429 aLserIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
701 TCTCGATTAAGAGACTTCTTCTGGAAGGAAATAAAGAACCGGGCGTTG 750
446 GlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe 462
751 GGTATCTGAAAACCTCATACCTTATATATTAAGATTAATTAATTAATATC 800
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
801 TCACCTTGGCACACACTGCTCGGATAGTCCACAGCCGCTCAACGACTTGG 850
479 aLserGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
851 TTAGCCAAAAAACAACTCAGCTGTCTGATATATACATCACGTTTAAATTC 900
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetGlu 512

```

901 GCATATGAAGCACTGAACCGTTTCATTCAGAAATATGATGATGATGCA 950
512 nargLeuLeuaspPthr 518
|||||
951 ACGTCTGCTAGATGACACG 969
seq_name: gb_ba2:AF167310

seq_documentation_block:
LOCUS AF167310 961 bp DNA BCT 17-APR-2000
DEFINITION Yersinia pestis strain Angola V antigen (lcrV) gene, complete cds.
ACCESSION AF167310
VERSION AF167310.1 GI:7578514
KEYWORDS Yersinia pestis.
SOURCE Yersinia pestis.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
REFERENCE 1 (bases 1 to 981)
AUTHORS Adair,D.M., Worham,P.L., Hill,K.K., Klevytska,A.M., Jackson,P.J., Friedlander,A.M. and Kelm,P.
TITLE Diversity in a variable-number tandem repeat from Yersinia pestis J. Clin. Microbiol. 38 (4), 1516-1519 (2000)
JOURNAL 20211685
REFERENCE 2 (bases 1 to 981)
AUTHORS Hill,K.K. and Jackson,P.J.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-1999) Bioscience Division, Los Alamos National Laboratory, MS:888, Los Alamos, NM 87545, USA
FEATURES
source
location/Qualifiers
1..981
/organism="Yersinia pestis"
/strain="Angola"
/db_xref="taxon:632"
1..981
/gene="lcrV"
1..981
/gene="lcrV"
/codon_start=1
/transl_table=1
/product="V antigen"
/protein_id="AAE64077.1"
/db_xref="GI:7578515"
/translation="MIRAYDONPHFIEDLENVNEVDLTGHSSVLELYVOLVYKNNIN
DISIKYPRKDEVFANRVITDIELRLKILAYLPEDALIKGHVYNOONGIKRVK
EELSSNTOMLEAFAMVAFHSTLADRIDDLIKVDSMMNHRGDARSKLRELALE
TAEIKTSVIOAEINKLSSGCTNINDEKSNLMDKNLYCTDEIFIRKASNEYILKEK
MPQTTIVDVSSEKTIYSIKDFGSEENRRTGALGNKLSYSTNKNNELSHFATTSQK
SRPNDIVDSQTTQLSDITSIRFNSAILELNRFKIQRKYSVQRLDDTSQK"
BASE COUNT 334 a 175 c 196 g 276 t
ORIGIN

alignment_scores:
Quality: 1628.00 Length: 326
Ratio: 5.009 Gaps: 0
Percent Similarity: 99.693 Percent Identity: 98.773

alignment_block:
US-08-699-716A-2 x AF167310 ..
Align seg 1/1 to: AF167310 from: 1 to: 981

196 MetLeuAlaTyrGluGlnAsnProGlnIHisPheIleGluAspLeu 212
|||||
1 ATGATTAGAGCCCTACGAAACAAACCCACACATTTATTATGAGATCTAGA 50
212 uysValArgValGluGlnLeuThrGlyHisGlySerSerValLeuGlu 229
|||||
51 AATGTTAGGCGTGAGAACACTTACGTGATGCTTCTTAGTTTNGAAG 100
229 luleValGlnLeuValLysAspLysAsnIleAspIleSerIleIleTyr 245
|||||

```

```

101 AATGGTTCAGTACGACAGATAAATAATAGATATTTCATTAAATAT 150
246 AspProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
151 GATCCAGAAAAGATCGAGGTTTGGCCATAGAGTAATTAATGATGA 200
262 PileGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThr 279
201 TATGAAATGCTCAGGAAAATCCCTAGCTTATTCTACCCAGAGATGCA 250
279 LeuLeuValGlyHisTyrAspAsnGluLeuGlnAsnGlyIleLysArg 295
251 TTCTTAAAGCGGCTCATTAATGACAACTGCAAAATGGCATTAAGCGA 300
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTrpGluLeuArg 312
301 GTAAGAAGTCTCTGAATCATCCCGAATACACAATGGGAATGGCGGC 350
312 AphenetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
351 GTTCATGGCAGTAAATGCATTTCTCTTAAACCGCGATCGATCATGATG 400
329 spLleuLeuValIleValAspSerMetAsnHisGlyAspAlaArg 345
401 ATATTTGAAAGTATGTTGATTCATCAATGAAATCATGATGATGCCGT 450
346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIle 362
451 AGCAAGTGGCGTGAAGAATAGCTGAGGTTACCCCGAATTAAGATTGA 500
362 rSerValIleGlnAlaGluIleAsnLysHisLeuSerSerGlyThr 379
501 TTCAGTTATTCACGCCGAAATTAATTAAGCATCTGTCTGTGTGTCACA 550
379 leaenIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
551 TAAATATCATGATTAATCATCAATATCATGATGATAAATTAATTAATGT 600
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeu 412
601 TATACAGATGAAGATTTTAAAGCCAGCCAGAGTCAAAATTTCTGCA 650
412 uLysMetProGlnThrThrIleGlnValAspGlySerGluLysIleVal 429
651 GAAATGCGCTCAAAACCACTTCAGTGTGATGGAGCGAGAAAAAATAG 700
429 alserIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
701 TCTGATTAAGGACTTCTTGGAGTGAATTAAGAAACCGGGCGCTTG 750
446 GlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeu 462
751 GGTATCTGAAAACCTACTACTATATAAAGATATAATATGATTAATTC 800
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeu 479
801 TCACCTTCCACACACACAGCTCGAATTAAGTCCAGCGCTCAACGACTGG 850
479 alserGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
851 TTAGCCAAAACAACTCAGCTGTGATTAATTAATCAACGTTTAAATCA 900
496 AlaIleGlnAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG 512
901 GCTATGAGACACTGAACGTTTCAATCAAGAAATATGATCAGATATGCA 950
512 nargLeuLeuAspAspThrSerGlyLys 521
951 ACGTCTGCTAGATGACAGCTGTGTAAG 978
seq_name: gb_dal:YPPPIVANT
seq_documentation_block:

```

```

LOCUS YPPPIVANT 981 bp DNA BCT 13-MAR-1997
DEFINITION Y.pseudotuberculosis V antigen gene.
ACCESSION X96802
VERSION X96802.1 GI:1405834
KEYWORDS V antigen.
SOURCE Yersinia pseudotuberculosis.
ORGANISM Yersinia pseudotuberculosis.
REFERENCE 1 (bases 1 to 981)
AUTHORS Roggenkamp, A., Geiger, A.M., Leitritz, L., Kessler, A. and
Heesemann, J.
TITLE Passive immunity to infection with Yersinia spp. mediated by
anti-recombinant V antigen is dependent on polymorphism of V
antigen
JOURNAL Infect. Immun. 65 (2), 446-451 (1997)
MEDLINE 97162308
REFERENCE 2 (bases 1 to 981)
AUTHORS Roggenkamp, A.
TITLE Direct Submision
JOURNAL Submitted (22-MAR-1996) A. Roggenkamp, Hygiene and Microbiology,
University of Wuerzburg, Josef-Schneider Strasse 2, 97080
Wuerzburg, FRG
FEATURES
source location/Qualifiers
1..981
/organism="Yersinia pseudotuberculosis"
/sub_species="type I"
/db_xref="taxon:633"
1..981
/codon_start=1
/transl_table=11
/product="V antigen"
/protein_id="CNA65594.1"
/db_xref="GI:1405835"
/db_xref="EMBL:U00001"
/translation="MIRAEONPQHFIDLEKRYVEQITGSSVLEEVLYVDKNI
DISIVDRKDSYEVANRVITDDLEKRIATVLPETALLKGHYDQQLQNGIKRYK
EFLESPNTOWELRAFMAVIFHSLADRIDDDILKVIYDSNNHGDANSKLRBLAEL
TAEKIVYIOAEIKNLHLSGGTINIHRKSNLMDKNYGYTDEIFEASAEYILEK
MQOTIOGETEERKIVSIKNFLESEKRRKGALGNLDSYNNKNNELSHPATGCSDK
SRPLNDIVSQKTLTQSDTITSFNSAIEALNFIQKIDSVMRLLDDTISGK"
BASE COUNT 341 a 174 c 191 g 275 t
ORIGIN
alignment_scores:
Quality: 1590.00 Length: 326
Ratio: 4.938 Gaps: 0
Percent Similarity: 98.773 Percent Identity: 96.626
Alignment block:
us-08-699-716a-2 x YPPPIVANT
Align seg 1/1 to: YPPPIVANT from: 1 to: 981
196 MetIleArgAlaTyrGluGlnAsnProGlnHisPheIleGluLysPleuG 212
1 ATGATTTAGAGCTTACGAAACAAACCAACATTTTATGAGATCTAGA 50
212 uLysValArgValGluGlnLeuThrGlyHisGlySerSerValLeuGluG 229
51 AAAAGTTAGGCTGACMACACTTACTGCTGCTTCTTCAGTTTGAAG 100
229 lueLeuValGlnLeuValLysAspLysAsnLleAspIleSerIleLysTyr 245
101 AATGGTTCAGTACGACAGATAAATAATAGATATTTCATTAAATAT 150
246 AspProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
151 GATCCAGAAAAGATTCGAGGTTTGGCCATAGAGTAATTAATGATGA 200
262 PileGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThr 279

```

201 TATCGAATGCTCAAGAAAATCCTAGCTTATTTCTACCCGAGATGCCA 250
 279 l e l e u l y s g l y h i s t y r a s p a n g l e u g l n a n g l y l e u s a r g 295
 251 TCTTAAGGCGGTCATTTATGACACCACTGCMAAATGGCATCAAGCA 300
 296 V a l l y s g l u p h e u g l u s e r s e r p r o a n t h r g l n t p g l u e u a r g l 312
 301 G T A A A G A G T T C C T T G A A T C A T C G C C G A T A C A A T G G A A T T G C G G C 350
 312 a p h e t a l a v a m e t h i s p h e s e r l e u t h r a l a s p a g l i e a s p a s a 329
 351 G T T C A T G C G A T A T A C A T T C T C T T T A C C C G G T C C T A T C A G A T A G 400
 329 s p i l e u l y s v a l l e v a l a s p s e r m e t a s n h i s t i s l y a s p a l a a r g 345
 401 A T A T T T G A A G T A G T A T G T T G A T C A T A T C A T C A T G T A T G C C C T 450
 346 S e r l y s l e u a r g l u g l u e u a l a g l u e u t h r a l a g l u e u l y s l e t y 362
 451 A C C A G T T C G T G A A G A A T T A C T G A C C G C C G A T T A A G A T T T A 500
 362 r s e r v a l l i e g l n a l a g l u i l e a n l y s h i s l e u s e r s e r s e r g l y t h r i 379
 501 T T C A G T A T T C A A C C G A A T T A A T A A G C A T C T G T A G T G C G C A C C A 550
 379 l a s n l l e h i s a p l y s s e r l e a s n l e u m e t a s p l y s a n l e u t y g l y 395
 551 T A A A T T C A T C A T A T A T C A T A T T A T C T A T G A T A A A A T T A T T A T T A T G 600
 396 T y r t h r a s p g l u g l u l e p h e l y s a l a s e r a l a g l u t y r l y l e u e u l 412
 601 T A T A C A G A T G A A G A G A T T T T A A G C C A G T G C A G A T A C A A A T T T C G A 650
 412 u l y s m e t p r o g l n t h r t h l l e g l n a l a s p g l y s e r g l u l y l y s l l e y 429
 651 G A A A A T G C C T C A A A C C A C C A T T C A G A A G G T G A G A C C G A A A A A A A T A T A G 700
 429 a l s e r l l e l y s a p h e l e u g l y s e r g l u a n l y s a r g t h r g l y a l a l e u 445
 701 T C T C A T A A A G A A C T T T C T T G A A A G T A G A A A A A A A A A A C C G G C G T T G 750
 446 g l y a n l e u l y s a n s e r t y r s e r t y r a n l y s a s p a n a n g l u e u s e 462
 751 G G T A A T C T G A A A G A C T A C T C T T A T A T A A G A T A A T A G A A T A T C 800
 462 t h i s p h e a l a t h r t h r c y s e r a s p l y s e r a r g p o l e u a n a s p l e u v 479
 801 T C A C T T T G C C A C C A C C T C C T C G A T A A G T C A G C C G C T C A A C A G A C T T G 850
 479 a l s e r g l n l y s t h r t h r g l n l e u s e r a s p l l e t h r s e r a r g p h e a n s e r 495
 851 T T A G C C A A A A A A C A C T A C G T G T C G A T T A T T A C A C A C G T T T A T T A T T C A 900
 496 A l a l e g l u a l e u a n a r p h e l l e g l n l y s t y r a s e r v a l m e t g l 512
 901 G G T A T T G A A G A C T A A C C G T T C A T T C A G A A A T A T A T A T C A G T A T G C A 950
 512 n a r g l e u l e u a s p a s p t h s e r g l y l y s 521
 951 A C G T C T G C T A G A T G A C A C G T C G T A A A 978

seq_name: gb_bal:YEPICRGVHP

seq_documentation_block:

LOCUS YEPICRGVHP 2201 bp DNA BCT 26-Apr-1993
 DEFINITION Yersinia pseudotuberculosis V-antigen (LcrG, LcrV, LcrH) genes,
 complete cds.
 ACCESSION M57893
 VERSION M57893.1 GI:155456
 KEYWORDS V-antigen,
 SOURCE Yersinia pseudotuberculosis (strain YPIII (PIBI) DNA.
 ORGANISM Yersinia pseudotuberculosis

REFERENCE
 1 (bases 1 to 2201)
 AUTHORS Bergman,T., Hakanson,S., Forsberg,A., Norlander,L., Macellaro,A.,
 Beckman,A., Boelln,I. and Wolf-Watz,H.
 TITLE Analysis of the V antigen lcrGH-yopBD Operon of Yersinia
 pseudotuberculosis: Evidence for a regulatory role of LcrH and LcrV
 JOURNAL J. Bacteriol. 173, 1607-1616 (1991)
 MEDLINE 91154114
 FEATURES
 source location/Qualifiers
 1..2201
 /organism="Yersinia pseudotuberculosis"
 /strain="YPIII (PIBI)"
 /db_xref="taxon:633"
 219..224
 /gene="lcr operon"
 219..258
 /gene="lcr operon"
 241..246
 /gene="lcr operon"
 252..258
 /gene="lcr operon"
 264..551
 /gene="lcrG"
 264..551
 /gene="lcrH"
 /codon_start=1
 /transl_table=1
 /evidence=experimental
 /product="V-antigen"
 /protein_id="AA27644.1"
 /db_xref="GI:155457"
 /translation="MKSSHFDKTLKQAEALANSDHRAKLQEMCDIGLTPAV
 MKFAQSAEELPAERLDELKREPOPHYDGRKPRFTMRGQIT"
 553..1533
 /gene="lcrV"
 553..1533
 /gene="lcrV"
 /codon_start=1
 /transl_table=1
 /evidence=experimental
 /product="V-antigen"
 /protein_id="AA27645.1"
 /db_xref="GI:155458"
 /translation="NIRAYBNQHRIEDLEKRVQLNGHSSYLELYQYKDKNI
 DISIKYDPKDSVFANRVITDDIELKILAFEDAILKGHYDNQKIKRVK
 EFLSSPNTQWELRAFMAYIHESLADRIDDDILKYIVSMNHGARSRLRELEL
 TALKIYSYQAEINRKLSSGGFINIDKRSINDKLNLYGTDEIFKASAEYKILEK
 MPOTIOEGTEKIKYSIKNLESEKRRGALGNLDSYSYNKDNNEISHFATCGSK
 SRPLNLYSGKTQLSDITSRFSALIEALNRFQKYDSVMGRLLDPTSGK"
 1546..2052
 /gene="lcrH"
 1546..2052
 /gene="lcrH"
 /codon_start=1
 /transl_table=1
 /evidence=experimental
 /product="V-antigen"
 /protein_id="AA27646.1"
 /db_xref="GI:155459"
 /translation="MOETPDQYOLAMESFLKGGGTIAMNEISSDLEOLYSIAF
 NOYOSKYEPAHKKVPOALCYLDHYDSRFELGACQANQOYIOLAHSSYSGAIINAK
 ERFPRFAECLLQKELDAESGLFLAOELLADKPEFELSTRVSSMLEAIKLKEM
 EHECVNPF"

BASE COUNT 714 a 450 c 485 g 552 t
 ORIGIN

alignment_scores:
 Quality: 1590.00 Length: 326
 Ratio: 4.938 Gaps: 0
 Percent Similarity: 98.773 Percent Identity: 96.626

alignment_block:
US-08-699-716a-2 x YEPLCRGHPH ..

Align seg 1/1 to: YEPLCRGHPH from: 1 to: 2201

```

196 MetIleArgAlaTyrGluInaSerProGlnHisPheIleGluAspLeuG 212
|||||
553 ATGATTAGACCTACGACAAACCAACCAATTTATTAAGGATCTAGA 602
|||||
212 ulysValArgValGluInaLeuThrGlyHisGlySerSerValLeuGluG 229
|||||
603 AAAGTAGGCTGGAACAACACTGCTACTGCTGCTCTCTGATTTTAAAG 652
|||||
229 IuLeuValGluLeuValLysAspLysAsnIleAspIleSerIleLysTyr 245
|||||
653 AATGTGCTCAGTAGTCAAAGATAAAAATATAGATATTTCCATTAATAT 702
|||||
246 AspProArgLysAspSerGluValPheAlaAsnArgValIleThrAspAs 262
|||||
703 GATCCCGAAAAGATTGCGAGGTTTGGCAATAGACTAATTTCTGATGTA 752
|||||
262 PileGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluAspThrI 279
|||||
753 TATCGATTCTCAGAAATCCAGCTTATTTCTACCCGAGATGCCA 802
|||||
279 IeLeuLysGlyGlyHisTyrAspAsnGluLeuGluAsnGlyIleLysArg 295
|||||
803 TTCTTAAGCGGTCATTATGACACCAACCAATGCGAAATGCGATCAGACGA 852
|||||
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTrpGluLeuArgAl 312
|||||
853 GTAAAGAGTTCCTTGAATCATCGCCCAATACAAAGGGAATTTGCGGCG 902
|||||
312 aPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAspA 329
|||||
903 GTTCATGCGCAGTATACATTCTCTTAACCGCGCATGATGATGATG 952
|||||
329 spIleLeuLysValIleValAspSerMetAsnHisGlyAspAlaArg 345
|||||
953 ATATTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1002
|||||
346 SerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTy 362
|||||
1003 AGCAAGTTCGCTGAGAGATTAAGCTGACCTTACCGCCCAATTAATAATTA 1052
|||||
362 rSerValIleGlnAlaGluIleAsnLysHisLysSerSerSerGlyThrI 379
|||||
1053 TTCAGTTATTCAGCTGAATTAATTAAGCATCTGTCGATGCGGCGCACCA 1102
|||||
379 IeAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
|||||
1103 TAAATATCCATGATTAATCAATTAATCTCATGATTAATAATTAATAGT 1152
|||||
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuG 412
|||||
1153 TATACGATTAAGAGATTTTAAAGCGACGACAGTACAAATTCGCA 1202
|||||
412 ulysMetProGlnThrThrIleGlnValAspGlySerGluLysLysIleV 429
|||||
1203 GAAATGCTCCCAACACCATTCAGGAAGGTGAGACCGAAAAAATAG 1252
|||||
429 aIleSerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
|||||
1253 TCTCGATTAAAGAACTTCTTGAAGAGTGAAGAAAAAGAACCGGCGCTTG 1302
|||||
446 GluAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnGluLeuLeu 462
|||||
1303 GGTAACTGAAAGACTCATCTCTTAATAAAGATTAATATGATTAATTC 1352
|||||
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
|||||
1353 TCACCTTGGCACACCGCTGCTGGATTAAGTCAGGCGCTCAAGACTTGG 1402

```

```

479 aIleSerLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
|||||
1403 TAGCCAAAAAACACTGCTGCTGATTAATACATGATTAATTAATCA 1452
|||||
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetI 512
|||||
1453 GCTATTAGACCTGAACCGTTTCATTCACAAATATGATTCAGTATGCA 1502
|||||
512 nArgLeuLeuAspAspThrSerGlyLys 521
|||||
1503 ACCTGCTGATGATGACACGCTGTGTTAA 1530
|||||
seq_name: gp_baz:AF102990
seq_documentation_block:
LOCUS AF102990 69673 bp DNA circular BCT 18-MAY-1999
DEFINITION Yersinia enterocolitica plasmid pYE227, complete sequence.
ACCESSION AF102990 AF054979 AF054980 AF054981 AF080156 Z69926 L06216
AF080155 AF022645 AF050104 AF054977 AF033863 U02499 U08019 U21297
M22781 U08222 Z18539 U94827
AF102990.1 GI:4324323
VERSION
KEYWORDS
SOURCE
ORGANISM
Yersinia enterocolitica.
Yersinia enterocolitica.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
1 (bases 32056 to 32871)
Cornelis, G., Sluiters, C., de Rouvoit, C.L. and Michiels, T.
Homology between virF, the transcriptional activator of the
Yersinia virulence regulon, and AraC, the Escherichia coli
arabinose operon regulator
J. Bacteriol. 171 (1), 254-262 (1989)
89123026
2 (bases 15074 to 15994)
Hakanson, S., Bergman, T., Vanooteghem, J.C., Cornelis, G. and
Wolf-Watz, H.
YopB and YopD constitute a novel class of Yersinia Yop proteins
Interact. Immun. 61 (1), 71-80 (1993)
93114907
3 (bases 45325 to 45717)
Mattiau, P. and Cornelis, G.R.
SydE, a chaparrone-like protein of Yersinia enterocolitica involved
in the secretion of YopE
Mol. Microbiol. 8 (1), 123-131 (1993)
93268087
4 (bases 23950 to 25269)
Woestyn, S., Allaoui, A., Mattiau, P. and Cornelis, G.R.
YscN, the putative energizer of the Yersinia Yop secretion
machinery
J. Bacteriol. 176 (6), 1561-1569 (1994)
94179088
5 (bases 29897 to 30961)
Allaoui, A., Woestyn, S., Sluiters, C. and Cornelis, G.R.
YscU, a Yersinia enterocolitica inner membrane protein involved in
Yop secretion
J. Bacteriol. 176 (15), 4534-4542 (1994)
94321323
6 (bases 43292 to 43717)
Mattiau, P., Bernier, B., Deslee, P., Michiels, T. and Cornelis, G.R.
Individual chaparrones required for Yop secretion by Yersinia
Proc. Natl. Acad. Sci. U.S.A. 91 (22), 10493-10497 (1994)
95024141
7 (bases 31537 to 31932)
Allaoui, A., Scheen, R., Lambert de Rouvoit, C. and Cornelis, G.R.
VirG, a Yersinia enterocolitica lipoprotein involved in Ca2+
dependency, is related to exsB of Pseudomonas aeruginosa
J. Bacteriol. 177 (15), 4230-4237 (1995)
95362644
8 (bases 53135 to 53497)
Stainier, I., Irlarte, M. and Cornelis, G.R.
YscM and YscM2, two Yersinia enterocolitica proteins causing
downregulation of yop transcription
Mol. Microbiol. 26 (4), 833-843 (1997)
JOURNAL

```

```

201 TATGGAATGCTCAGAAAATCCTAGCTATTCTTACCCGAGATGCCA 250
279 leuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
251 TTCTTAAGGCGGCTATTATGACACCACTGCAAAATGGCATACACGA 300
296 ValLysGlnPheLeuGlnSerSerProAsnThrGlnTrpGluLeuArgAl 312
301 GTAAAGAGTTCCTTGATCATCGCCAAATACACATGGGAATGGCGGC 350
312 apHeMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
351 GTTCATGCGAGTATACATCTCTTACCGCGCATGATCATGATG 400
329 spLleLeuLysValIleValAspSerMetAsnHisHisGlyAspAlaArg 345
401 ATATTGAAAGATGATGTGATCAATCAATCAATGATGATGCCCGT 450
346 SerLysLeuArgGlnGluLeuAlaGluLeuThrAlaGluLeuLysIleTy 362
451 AGCAAGTGGGTGAGAAATTAAGTACGCTTACCGCCGATTAAGATTTA 500
362 rSerValIleGlnAlaGluIleAsnLysHisLeuSerSerSerGlyThrI 379
501 TTCAGTATTCACACCGCAAAATTAATAGCATGCTGATAGTGGCACCA 550
379 leaAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
551 TAAATATCCAGTAATCAATATCATGATGATTAATAATTAATATAGGT 600
396 TyrThrAspGlnGluIlePheLysAlaSerAlaGluTyrLysIleLeuG1 412
601 TATACAGATGAGAGATTTTAAAGCCAGTGCAGATCAAAATTCGCA 650
412 uLysMetProGlnThrThrIleGlnValAspGlySerGluLysLysIleV 429
651 GAAATGCTTCAACCAACCATTCAGAGAGTGAAGACGCAAAAAAATAG 700
429 alSerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
701 TCTCGATTAAGAACCTTCTTGAAGTGAAGAAAAAAGAACCGGCGCTTG 750
446 GlnValAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe 462
751 GGTATCTGAAAGACTACTACTTATATAAAGATAATTAATGAATTTC 800
462 rHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
801 TCACTTGGCACACAGCTGCTCGATTAAGTCCAGGCGCTCAACGACTTG 850
479 alSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
851 TTACCCAAAAAACAACCTGCTGATTAATCACTACGCTTTTAATCA 900
496 AlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG1 512
901 GCTATTGAGACACTGACCGTTTCATTGAGAAATATGATTCAGTATGCA 950
512 narGluLeuAspAspThrSerGlyLys 521
951 ACGTCTGCTAGATGACAGCTCTGGTAA 978
seq.name: gb_dal.YEPLCRGVP
seq_documentation_block:
LOCUS YEPLCRGVP 2201 bp DNA BCT 26-APR-1993
DEFINITION Yersinia pseudotuberculosis V-antigen (LcrG, LcrV, LcrH) genes,
complete cds.
ACCESSION M57893
VERSION M57893.1 GI:155456
KEYWORDS V-antigen.
SOURCE Yersinia pseudotuberculosis (strain YPIII (PIB1) DNA.
ORGANISM Yersinia pseudotuberculosis

```

```

REFERENCE 1 (bases 1 to 2201)
AUTHORS Bergman,T., Hakansson,S., Forsberg,A., Norlander,L., Macellaro,A.,
Baekman,A., Boelin,I. and Wolf-Watz,H.
TITLE Analysis of the V antigen lcrGVH-yopD Operon of Yersinia
pseudotuberculosis: Evidence for a regulatory role of LcrH and LcrV
J. Bacteriol. 173, 1607-1616 (1991)
JOURNAL MEDLINE 91154114
FEATURES
source location/Qualifiers
1..2201 /organism="Yersinia pseudotuberculosis"
/strain="YPIII(PIB1)"
/db_xref="taxon:633"
-35_signal 219..224
gene 219..258
-10_signal 241..246
RBS 252..258
gene 264..351
CDS 264..551
/codon_start=1
/transl_table=11
/evidence=experimental
/product="V-antigen"
/protein_id="AAA27644.1"
/db_xref="GI:155457"
/translation="MKSSHPEYDTLTKQAEILADSPHRAKLQEMADIGLPEAV
MTFNGRAHEIKPARELIDIKRERQPHDPGRPRKPTMARQIT"
553..1533
/gene="lcrV"
553..1533
CDS 553..1533
/gene="lcrV"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="V-antigen"
/protein_id="AAA27645.1"
/db_xref="GI:155458"
/translation="MIRAEQNPQHFIEDLEVRYEQLTGHSSVLEELVQYVKNTI
DISIKYDKDSVFRANRYTDDIELKRLVFLPEDAILGGHYDNLQNGIRVK
EFLSEPTQWELRAFMAVTHFSLNADRIDDLIVYDSNMHBDARKLEIAEL
TAEKIVSVIOAEIKRHLSSGCTINIDKSLMDKNLYGDEEIFKASAEYKILEL
MNOTIOEGETEKKIVSIKNFLESKKRKGALGNKDSYSYKDNNEISHFATCSDK
SRPLDIVSOKTQQLSDITSRNSAIEALNRIQKYDSVMQRLDDTSK"
1546..2052
/gene="lcrH"
1546..2052
CDS 1546..2052
/gene="lcrH"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="V-antigen"
/protein_id="AAA27646.1"
/db_xref="GI:155459"
/translation="MOERTDQEOVLAMESFLKGGTITAMLEISDTEQLYSIAF
NOYSGKYEDAHKVFQALCVLDHYDSRFLGACROAGQVDLAIHSYGAINDIK
ERFPFHAECILQKGEIAEASGLFLAODELTADREFRELESTRVSSMLEAIKKEM
EHECVDP"
BASE COUNT 714 a 450 c 485 g 552 t
ORIGIN
alignment_scores:
Quality: 1590.00 Length: 326
Ratio: 4.938 Gaps: 0
Percent Similarity: 98.773 Percent Identity: 96.626

```


REFERENCE	AUTHORS	TITLE
JOURNAL MEDLINE 98087284	9 (bases 18681 to 18968)	
JOURNAL MEDLINE 98143428	Boyd,A.P., Sory,M.P., Iriarte,M. and Cornelis,G.R.	Heparin interferes with translocation of Yop proteins into HeLa cells and binds to LcrG, a regulatory component of the Yersinia yop apparatus
JOURNAL MEDLINE 98190073	Mol. Microbiol. 27 (2), 425-436 (1998)	
JOURNAL MEDLINE 98190073	10 (bases 22612 to 22890)	
JOURNAL MEDLINE 98190073	Iriarte,M., Sory,M.P., Boland,A., Boyd,A.P., Mills,S.D., Lamberton,I. and Cornelis,G.R.	YjveA, a protein involved in control of Yop release and in translocation of Yersinia Yop effectors
JOURNAL MEDLINE 98190073	EMBO J. 17 (7), 1907-1918 (1998)	
JOURNAL MEDLINE 98190073	11 (bases 5 to 69673)	
JOURNAL MEDLINE 98190073	Iriarte,M., Lamberton,I., Kerbourch,C. and Cornelis,G.R.	Detailed genetic map of the pYve227 plasmid of Yersinia enterocolitica serotype O:9
JOURNAL MEDLINE 98190073	12 (bases 1 to 69673)	
JOURNAL MEDLINE 98190073	Iriarte,M., Lamberton,I., Kerbourch,C. and Cornelis,G.R.	Direct Submission
JOURNAL MEDLINE 98190073	Submitted (30-OCT-1998), Microbial Pathogenesis Unit, Christian de Duve Institute of Cellular Pathology and Faculte de Medecine, Universite Catholique de Louvain, Av. Hippocrate 74,49, Brussels 1200, Belgium	
JOURNAL MEDLINE 98190073	On May 18, 1999 this sequence version replaced g1:3411136	
JOURNAL MEDLINE 98190073	g1:3411135 g1:3411139 g1:3411137 g1:3603360 g1:1524380 g1:15536	
JOURNAL MEDLINE 98190073	g1:3603357 g1:2738986 g1:2944089 g1:3411138 g1:3108216 g1:437201	
JOURNAL MEDLINE 98190073	g1:497215 g1:885971 g1:155526 g1:483322 g1:48593 g1:2735421.	
JOURNAL MEDLINE 98190073	Location/Qualifiers	
JOURNAL MEDLINE 98190073	1. .69673	
JOURNAL MEDLINE 98190073	/organism="Yersinia enterocolitica"	
JOURNAL MEDLINE 98190073	/strain="W22703"	
JOURNAL MEDLINE 98190073	/serotype="O:9"	
JOURNAL MEDLINE 98190073	/db_xref="taxon:630"	
JOURNAL MEDLINE 98190073	/plasmid="pYve227"	
JOURNAL MEDLINE 98190073	53. .203	
JOURNAL MEDLINE 98190073	complement(889. .1434)	
JOURNAL MEDLINE 98190073	/note="ORF181"	
JOURNAL MEDLINE 98190073	/codon_start=1	
JOURNAL MEDLINE 98190073	/transl_table=1	
JOURNAL MEDLINE 98190073	/product="unknown"	
JOURNAL MEDLINE 98190073	/protein_id="AAD16801.1"	
JOURNAL MEDLINE 98190073	/db_xref="gi:4324324"	
JOURNAL MEDLINE 98190073	/translation="MARVYVGQYGIKFNKNGLDSPAIQRCGLVLAISDQHYLCA	
JOURNAL MEDLINE 98190073	HPDSAFGLQNLQDIREIENGIRMSLRATYVGSGKQSYLRCSPPSHIGELIYN	
JOURNAL MEDLINE 98190073	FMKSHGALQYNSQYISVFPETNFRHOKCGCITKQNPFDLGGSPQAMNLAQRTI	
JOURNAL MEDLINE 98190073	KLRSEYSLTHAKMLDISRLT"	
JOURNAL MEDLINE 98190073	1616. .2083	
JOURNAL MEDLINE 98190073	/note="ORF155"	
JOURNAL MEDLINE 98190073	/codon_start=1	
JOURNAL MEDLINE 98190073	/transl_table=1	
JOURNAL MEDLINE 98190073	/product="unknown"	
JOURNAL MEDLINE 98190073	/protein_id="AAD16802.1"	
JOURNAL MEDLINE 98190073	/db_xref="gi:4324325"	
JOURNAL MEDLINE 98190073	/translation="MINTFPELLQKIASHGDLKLSODEYGLCELLINDRVVYMLRA	
JOURNAL MEDLINE 98190073	DEINRLTLLPILGFGSPERSTASQLEFCYSINALNKQDPCAWSEEGELTAFKRL	
JOURNAL MEDLINE 98190073	SLGLANVSKELANFYDMLSLVSLPRTQDELPHRTQSIQSTQSYQVAKG"	
JOURNAL MEDLINE 98190073	2091. .4280	
JOURNAL MEDLINE 98190073	/gene="yopO"	
JOURNAL MEDLINE 98190073	2091. .4280	
JOURNAL MEDLINE 98190073	/gene="yopO"	
JOURNAL MEDLINE 98190073	/codon_start=1	
JOURNAL MEDLINE 98190073	/transl_table=1	
JOURNAL MEDLINE 98190073	/product="protein kinase YopO"	
JOURNAL MEDLINE 98190073	/protein_id="AAD16803.1"	
JOURNAL MEDLINE 98190073	/db_xref="gi:4324326"	
JOURNAL MEDLINE 98190073	/translation="MKIMGIMPISLAKAHERISQHWNPVGLNIGKRYRIINDNQ	
JOURNAL MEDLINE 98190073	VLRLNAGFSLPREGVQKIFSGKMFNSIARNLTETLHAQKTSQSLSDIPNAL	
JOURNAL MEDLINE 98190073	NLFAKQPTPELPICMGKGPLSGAPDIEGRNRAETDKPAEGSHSITILETDNQYLAK	
JOURNAL MEDLINE 98190073	IERSTAECHLFAELPAIKHYTKIKGKHNLANVGMAYPIGNKEEDLLADEVDGR	

alignment_scores: 1557.00
Quality: 4.896
Ratio: 4.896

length: 323
gaps: 0

Percent Similarity: -98.452 Percent Identity: 95.356

Alignment block:

US-08-699-716A-2 x AF102990/rev ..

Align seg 1/1 to reverse of: AF102990 from: 1 to: 69673

```

196 MetIleatgaIaTyGluGlnAsnProGlnHsPheIleGluAspLeuG1 212
|||||
18679 ATGATTAGAGCTTACGACAAACCCACACATTATTATGAGATCTGA 18630
|||||
212 uLysValArgValGluGlnLeuThrGlyHsGlySerSerValLeuG1 229
|||||
18629 AAAAGTTGGGTGGACACACTTACTGTCATGCTTCCTCAGTTTAGAG 18580
|||||
229 LuLeuValGlnLeuValLysAspLysAsnIleAspIleSerIleLysTyr 245
|||||
18579 AATTGGTTCAGTTAGTCAAGATAAAGATGATATTCCATTAAATAT 18530
|||||
246 AsProAlaGlyAspSerGluValAlaPheAlaAsnArgValIleThrAspAs 262
|||||
18529 GATCCCAAAAAGATTCGGAAGTTTTCGCCGATAGGTAATTACTGATGA 18480
|||||
262 PIIleGluLeuLysLysIleLeuAlaTyrPheLeuProGluAspThrI 279
|||||
18479 TATCGAATTACTCAGAAATCTCTGCTTATTTTACCCTGAGATGCCA 18430
|||||
279 LeLeuLysGlyGlyHsIleTyrAspAsnGlnLeuGlnAsnGlyIleLysArg 295
|||||
18429 TTCTTAAAGCGCGGTATATGACAAACCACTGCAAAATGGCATCAAGCGA 18380
|||||
296 ValLysGluPheLeuGluSerSerProAsnThrGlnTpgIuLeuArgAl 312
|||||
18379 GTAAAGAGTTCCTGGAATCATCGCCGATACACATGGAGTTGCGGCG 18330
|||||
312 aPheMetAlaValMetHsPheSerLeuThrAlaAspArgIleAspAspA 329
|||||
18329 GTTCATGCGCAATATGCAATTTCTTCTTAAACCCGATCGTATCGATGAG 18280
|||||
329 sPIleLeuLysValIleValAspSerMetAsnHsIleGlyAspAlaArg 345
|||||
18279 ATATTGTGAAGTGATGTGTGATCATGATCATCATGATGATGCCGT 18230
|||||
346 SerLysLeuArgGluGluLeuAlaIuLeuThrAlaGluLeuLysIleTy 362
|||||
18229 GGCAAGTTCGCTGAAGATTAGCTAGCTTACCCCGAATTAAAGATTGA 18180
|||||
362 rSerValIleGlnAlaGluIleAsnLysHsIleuSerSerSerGlyThrI 379
|||||
18179 TTCAGTTATTCAGCCGAATTAATAGCATCTGCTTAATAGTGACACCA 18130
|||||
379 leAsnIleHsAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGly 395
|||||
18129 TAAATATCATCATGATTAATCATATTCATCATGATGATGATTAATATG 18080
|||||
396 TyrThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuG1 412
|||||
18079 TATACAGATGAGAGATTTTAAACCCAGCGAGATACAAAAATTCCTCGA 18030
|||||
412 uLysMetProGlnThrIleGlnValAspLysSerGluLysLysIleY 429
|||||
18029 GAAAATGCTCAAAACCATTAAGGAAGGTGAGACCGAAAAAAAATATG 17980
|||||
429 aLSerIleLysAspPheLeuGlySerGluAsnLysArgThrGlyAlaLeu 445
|||||
17979 TCTCGATAAGACCTTCTTGAAAGTGAGAAATAAAGAACCGGGCGTTG 17930
|||||
446 GlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSe 462
|||||
17929 GGTAACTCGAAGATTCATCTATATAATAAAGATAATGAAATATATC 17880
|||||
462 rHsPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuV 479
|||||

```

```

17879 TCACCTTCCACACACCTGCTCGGATTAAGTCCAGGCCGCTCACACACTTGG 17830
479 aLSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSer 495
|||||
17829 TTAGCCAAAAAACAACTCAGCTGCTGATATATACATCACGTTTAATTCA 17780
|||||
496 AlaIleGlnAlaLeuAsnArgPheIleGlnLysTyrAspSerValMetG1 512
|||||
17779 GCTATGAGACACTGAAACCGTTTCATTCAGAAATATGACTCAGTATGCA 17730
|||||
512 nArgLeuLeuAspAspThr 518
|||||
17729 ACGTCTGCTAGATGACAG 17711

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

56 hrLeuValGluProAlaArgIleThrLeuThrTyrIleGluValAlaPro 72
 110 CTCTGTTGAACCGCCGATCCTTACATATATAGAGAGCGCTCA 159
 73 IleThrIleMetAspAsnGlyAsnIleAspThrGluLeuLeuValGlyThr 89
 160 ATTACATATATGACATGGAATGGAATCATGATACGAATTTCTGTTGTCAC 209
 89 rLeuThrLeuGlyGlyTyrTyrLeuThrGlyThrThrSerThrSerValAsp 106
 210 GCTTACTCTTGGCGGTATATAAAGAGAACCATACATCATCTGTTAACT 259
 106 heThrAspAlaIleGlyAspProMetTyrLeuThrPheThrSerGlnAsp 122
 260 TTACAGATGCCGGGTGATCCATGCTACTTAACATTTCTCTCGGAT 309
 123 GlyAsnAsnHISGlnPheThrThrLeuValIleGlyAspSerArgAs 139
 310 GGAATATACCCACCATTCCTACATAAAGATGGCAGAGATTCTAGACA 359
 139 pPheAspIleSerProLysValAsnGlyLysLeuValGlyAspAsp 156
 360 TTTGATATCTCTCTTAAGTAAACGCTGAGAACTTGTGGGGATGACG 409
 156 aIValLeuAlaThrGlySerGlnAspPheValArgSerIleGlySer 172
 410 TCGTTTGGCTACGGGCGACCGAGATTCTTGTTCCTCAATGGTTC 459
 173 LysGlyGlyLysLeuAlaIleGlyLysTyrThrAspAlaValThrValTh 189
 460 AAGCGGTAACCTTCACACAGTAATACACTGATGCTGAACCGTAAC 509
 189 rValSerAsnGln.....GluPheMetIleArgAlaTyrGluGlnA 203
 510 CTATATCAACCAAGATCCATCGATGCTGATTAAGCTTACGCTAGAACAA 559
 203 snProGlnHisPheIleGluAspLeuGluLysValArgValGluGlnLeu 219
 560 ACCCAACAACTTTTATGAGATCTAGAAAAGTTAGGTGGAACAACCT 609
 220 ThrGlyHisGlySerSerValLeuGluGluLeuValGluLeuValLysAs 236
 610 ACTGTCATGCTTCTTCACTTTAGAGATGCTTCACTTACTGTAACA 659
 236 pLysAsnIleAspIleSerIleLysTyrAspProArgLysAspSerGly 253
 660 TAAATATATAGATATTTCCATTAATATATGATCCAGAAAAGATTGCGAG 709
 253 aIPheAlaAsnArgValIleThrAspAspIleGluLeuLeuLysIle 269
 710 TTTTGGCCATAGAGTAATTAAGATGATGAAATGCTCAGAAAATC 759
 270 LeuAlaTyrPheLeuProGluAspThrIleLeuLysGlyIleHisTyrAs 286
 760 CTAGCTTATTTCTACCCAGAGATGCAATCTTAAAGCGGCTCATATATA 809
 286 pAsnGluLeuGlnAsnGlyIleLysArgValLysGluPheLeuGlnSer 303
 810 CAACCAACTGCAAAATGCGATCAAGGAGTAAAGAGTCCCTGGAATCAT 859
 303 exProAsnThrGlnTrpGluLeuArgAlaPheMetAlaValMetHisPhe 319
 860 CGCCCAATACACAAATGGGAATGGCGGCTCATGCAATAATGCAATTC 909
 320 SerLeuThrAlaAspArgIleAspAspAspIleLeuLysValIleValAs 336
 910 TCTTAAACCGCGATCGTATCATGATGATATTTGAAAGTATGTTTGA 959
 336 pSerMetAsnHisGlyAspAlaArgSerLysLeuArgGluGluLeu 353
 960 TTCATATGATCATCATGATGATGCGGTGCAAGTTGCTGAGAGATATG 1009
 353 IaGluLeuThrAlaGluLeuLysIleTyrSerValIleGlnAlaGluIle 369

1010 CTGAGCTTACCGCCGAATTAAGATTATTCAGTTTCAAGCCCAAT 1059
 370 AsnLysHisLeuSerSerSerGlyThrIleAsnIleHisAspLysSerI 386
 1060 AATAAGCATCTCTTAGTAGTGTGCGACCATTAATATCCATGATGATCAT 1109
 386 eaSnLeuMetAspLysAsnLeuTyrGlyTyrThrAspGluIlePheL 403
 1110 TAATCTCATGATTAATAAATTTATATGTTATACAGATGAAGATTTTTA 1159
 403 yAlaSerAlaGluTyrLysIleLeuGlyLysMetProGlnThrThrIle 419
 1160 AAGCCAGCGCAGATGACAAATTCGAGAAAATGCTCAACCCACCAT 1209
 420 GlnValAspGlySerGlyLysLysIleValSerIleLysAspPheLeuG 436
 1210 CAGGTGATGGAGCGGAGAAAAAATAGTCTCGATTAAGGACTTCTTGG 1259
 436 ySerGluAsnLysArgThrGlyAlaLeuGlyAsnLeuLysAsnSerTyr 453
 1260 AAGTGAATATAAAGAACCGGCGCTGGTATCTGAATAAATCATACT 1309
 453 eTyrAsnLysAspAsnAsnGluLeuSerHisPheAlaThrThrCysSer 469
 1310 CTATATAATAAAGATATATGATATATCTACTTGGCCACACCGCTGCG 1359
 470 AspLysSerArgProLeuAsnAspLeuValSerGlnLysThrThrGlnLe 486
 1360 GATAGTCCAGCGCCCTCAGACACTTGTGTAACCAAAAACACTCACT 1409
 486 uSerAspIleThrSerArgPheAsnSerAlaIleGluAlaLeuAsnArg 503
 1410 GTCTATATTAATCATCACTTTAATTCAGCTATTAACACTGAACCGCT 1459
 503 heIleGlnLysTyrAspSerValMetGlnArgLeuLeuAspAspPheSer 519
 1460 TCATTCAGAAATATGATTCAGTATGATGCAACGTCCTGATGACACGCT 1509
 520 GlyLys 521
 1510 GGTAAA 1515
 seq_name: N_Geneseq_36:T38256
 seq_documentation_block:
 ID T38256 standard; DNA; 1462 BP.
 AC T38256;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1/V antigen gene fusion.
 KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
 RM F1 antigen; cagl; ds.
 OS Chimeric Yersinia pestis strain GB;
 FH key Location/Qualifiers
 FT cds 8..1450
 FT /*tag= a
 FT /product= F1/V fusion protein
 FT 452..472
 FT /*tag= b
 FT /note= "bases 452-472 is a sequence derived
 from PCR primers"
 FT misc_feature
 PN WO9628551-A1.
 PD 19-SEP-1996.
 PE 13-MAR-1996: G00571.
 PR 13-MAR-1995: GB-005059.
 PR 13-SEP-1995: GB-018946.
 PR 05-DEC-1995: GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RM, Williamson ED;
 DR WPI: 96-433824/43.
 DR P-PSDB; W01045.
 Yersinia pestis V antigen and F1 antigen or their protective

DT 18-APR-1996 (first entry)
 DE Partial LcrV (V antigen) gene of Y. pestis.
 KW LcrV; V antigen; virulence; plague; vaccine; epitope; ss.
 OS Yersinia pestis.
 FH key
 FT cds Location/Qualifiers
 FT 1..990
 FT /tag- a
 FT /note- "V antigen"
 PN MO9524475-A1.
 PD 14-SEP-1995.
 PE 06-MAR-1995; G00481.
 PR 08-MAR-1994; GB-004577.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Leary SEC, Tiltball RW, Williamson ED, Leary SE;
 DR WPI: 95-328268/42.
 DR P-PSDB: R79961.
 PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
 PT oral or parenteral vaccines for protection against plague
 PS claim 6: Page 11-13: 25pp: English.
 CC T04222-23 are DNA sequences (LcrV) encoding all or a protective epitopic
 CC part of the mature V protein of Yersinia pestis. The protein was
 CC expressed as a fusion protein with maltose binding protein or
 CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis is
 CC the highly virulent causative organism of plague in a wide range of
 CC animals, including man. The V antigen (LcrV) is an unstable 37.3 kDa
 CC monomeric peptide encoded on the ca. 70 kb Lcr plasmid. The V antigen
 CC is postulated to act as a virulence antigen, and transformed
 CC microorganisms contg. recombinant DNA encoding a V antigen protein/
 CC peptide are useful in vaccines to protect against plague.
 SQ Sequence 1014 BP; 346 A; 181 C; 201 G; 286 T;

alignment_scores:
 Quality: 1641.50 Length: 328
 Ratio: 5.020 Gaps: 1
 Percent Similarity: 99.695 Percent Identity: 99.390

alignment_block:
 US-08-699-716A-2 x T04222 ..

Align seg 1/1 to: T04222 from: 1 to: 1014

194 GluPhmetIleArgAlaTyrGlnAsnProGlnHisPheIleGluAs 210
 7 GAATTC...ATTAGAGCCTACGACAAACCCACACATTTATTAGAGA 53
 210 pLeuGluLysValArgValGluGlnLeuThrGlyHisGlySerValL 227
 54 TCTGAAAAAATTAGGTGGTGAACAACCTACGTGTCATGCTTCTTACGTT 103
 227 eugGluLeuValGlnLeuValLysAspLysAsnIleAspIleSerIle 243
 104 TAGAAGAAATTGCTTCAGTACGAAAGATAAAATATGATATTCCATT 153
 244 LysTyrAspProArgLysAspSerGluValPheAlaAsnArgValIleTh 260
 154 AAAATATGATCCCAAGAAAGATTCGAGGTTTTCGCAATAGAGTATAC 203
 260 rAspAspIleGluLeuLeuLysLysIleLeuAlaTyrPheLeuProGluA 277
 204 TGATGATATCGAATTCCTCAGAAATCTTAGCTATTATTCTACCCGAG 253
 277 sPthrIleLeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyIle 293
 254 ATGCATATCTTAAAGCGCGTCATTTGACAAACCACTGCAAAATGCATC 303
 294 LysArgValLysGluPheLeuLysSerSerProAsnThrGlnTyrGluLe 310
 304 AAGCGAGTAAAGAGTCTCTGAATCATCGCGGAATACACATGGGAATT 353
 310 uArgAlaPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleA 327
 354 GCGGCGCTTCATGCGAGTATGATCATTTCTTTAAACCGCCGATCTATCG 403

327 sPaSPaSPiLeuLysValIleValAspSerMetAsnHisGlyAsp 343
 404 ATGATGATATTTTGAAGTGAATGCTGTGATTCATGAATGATCATCATGATG 453
 344 AlaArgSerLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLys 360
 454 GCGCGTACCAAGTTGCGTGAAGAAATTAAGCTGAGCTTACCGCGCAATTAA 503
 360 sIleTyrSerValIleGlnAlaGluIleAsnLysHisLeuSerSerGerg 377
 504 GATTATTCAGTATTTCACGCCGAATTAATTAACATCTGCTAGTGAGTG 553
 377 LysThrIleAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeu 393
 554 GCACCATTAATATCCATGATTAATCCATTAATCTCATGAGATTAATAATTTA 603
 394 TyrGlyTyrThrAspGluGluIlePheLysAlaSerAlaGlyTyrLysII 410
 604 TATGCTTATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAAT 653
 410 eLeuGluLysMetProGlnThrThrIleGlnValAspGlySerGluTyrSL 427
 654 TCTGAGAAATGCTCTCAAAACCATTCAGCTGATGCGGAGCAGAGAAA 703
 427 yslIleValSerIleLysAspPheLeuGlySerGluAsnLysArgThrGly 443
 704 AAATAGTCTCGATTAAGAGCTTCTTGGAAGTGAAGATTAAGAACCGGG 753
 444 AlaLeuGlyAsnLeuLysAsnSerTyrSerTyrAsnLysAspAsnGlnI 460
 754 GCGTGGGTAATCTGAANAACCTACACTTATATATTAAGTAAATATATGA 803
 460 uLeuSerHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnA 477
 804 ATTATCTCAGCTTTCACACACCTGCTCGGATTAATCCAGCGCGCTCAAG 853
 477 sPheValSerGlnLysThrThrGlnLeuSerAspIleThrSerArgPhe 493
 854 ACTTGCTTACCCAAAAAACAACCTCAGCTGTCTGATATTAATCATCAGTTT 903
 494 AsnSerAlaIleGluAlaLeuAsnArgPheIleGlnLysTyrAspSerVa 510
 904 AATTCAGTATTGAAGCAGTGAACCGTTTCATTTCAGAAATATGATTCAGT 953
 510 lMetGlnArgLeuLeuAspAspThrSerGlyLys 521
 954 GATGCAACGTCTGCTAGATGACAGCTCTGCTGTA 987
 seq_name: N_Geneseq_36:T38242
 seq_documentation_block:
 ID T38242 standard; DNA: 1014 BP.
 AC T38242;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis LcrV (V antigen) gene.
 KW Plague; vaccine; genetic immunisation; V antigen; LcrV;
 KW FI antigen; ds.
 OS Yersinia pestis strain GB.
 FH key
 FT cds Location/Qualifiers
 FT 1..990
 FT /tag- a
 FT /tag- a
 FT /tag- b
 FT /note- "vector pMAL-p2 or pMAL-c2-derived bases"
 FT mutation
 FT 1006
 FT /tag- c
 FT /note- "base 1006 is altered to a T to create a
 second in-frame stop codon"
 PN MO9628551-A1.
 PD 19-SEP-1996.
 PE 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.

PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MTNA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
 DR WPI: 96-433824/43.
 DR P-PSD: W01040.
 FT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Claim 5; Page 25-28; 98PP; English.
 CC A lcrv gene sequence (T38242) codes for the Yersinia pestis V
 CC antigen (W01040), which is capable of evoking protective immune
 CC responses in animals. The gene was amplified from Y. pestis
 CC DNA by PCR using primers (T38250-51) homologous to the 5' and 3'
 CC ends of the gene. The gene was inserted into vector pMAL-p2,
 CC pMAL-c2 or pGEX-5x-2 (see also T38243) to allow prodn. of
 CC recombinant V antigen for use in vaccines against plague.
 CC Expression in gut-colonising organisms and attenuated Salmonella
 CC typhi allows prodn. of live vaccines. F1/V antigen fusions were
 CC also created (see also T38249 and T38256). The gene can itself be
 CC used in genetic vaccines.
 SQ Sequence 1014 BP; 347 A; 180 C; 201 G; 286 T;

alignment_scores:
 Quality: 1641.50 Length: 328
 Ratio: 5.020 Gaps: 1
 Percent Similarity: 99.695 Percent Identity: 99.390

alignment_block:
 US-08-699-716a-2 x T38242 ..

Align seg 1/1 to: T38242 from: 1 to: 1014

194 GluphemeTlleRgAlaTyrgLugInuAsnProGlnHisPheIleGluAs 210
 |||||
 7 GAATTC..ATTAGACCTACGACAAACCCACACATTTATATAGGA 53
 210 pleugluTyValArgValGluInuLeuThrgLysGlySerSerVal 227
 |||||
 54 TCTAGAAAAGTAGGGTGAACAACCTTACTGTGTCATGTTCTCAGTTT 103
 227 eugluGluLeuValGlnLeuValLysAspLysAsnIleAspIleSerIle 243
 |||||
 104 TAGAAGAATTGGTTCACTAGTCAAAAGATATAATATGATATTCACAT 153
 244 LysTyAspProArgLysAspSerGluValPheAlaAsnArgValIleTh 260
 |||||
 154 AATATATGATCCCGAAGAAAGATTGGAGGTTTGGCAATATAGATATATTC 203
 260 rAspAspIleGluLeuLeuLysLysIleLeuAlaTyRheLeuProGluA 277
 |||||
 204 TGTATGATATCGAATTCCTCAAGAAAATCCTAGCTTATTTCTACCGAGG 253
 277 spThrIleLeuLysGlyGlyHisTyAspAsnGlnLeuGlnAsnGlyIle 293
 |||||
 254 AAGCCATCTCTAAAGCGGTCATTATGACAACTGCAAAATGGAATC 303
 294 LysArgValLysGluPheLeuGluSerSerProAsnThGlnTrpGluLe 310
 |||||
 304 AAGCGGTAAGAGTTCCTGATCATCGCCGAATACAAAGGGAATTC 353
 310 uArgAlaPheMetAlaValMetHisPheSerLeuThrAlaAspArgIleA 327
 |||||
 354 GCGGGGTCATGCGAGTAATGATTCCTTTAACCAGCCGAGATCATGCG 403
 327 spAspAspIleLeuLysValIleValAspSerMetAsnHisGlyAsp 343
 |||||
 404 ATGATGATATTTTGAAGATGATGTGATCATGATCATGATCATGATGAT 453
 344 AlaArgSerLysLeuArgGluGluLeuAlaGluLeuThrAlaLeuLeu 360
 |||||
 454 GCCCGTAGCAAGTTGGTGAAGATTAAGCTGACCTTACCGCCGAATTTAA 503

360 sIleTySerValIleGlnAlaGluIleAsnLysHisLeuSerSerSerg 377
 |||||
 504 GATTATTCAGTATTATCAAGCCGAATTAATATGACATCTGCTACTAGTG 553
 377 LyThrIleAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeu 393
 |||||
 554 GCACATTAATATTCATGATTAATCACTTAATCTGATGATTAATAATTTA 603
 394 TyrgLyTyThrAspGluGluIlePheLysAlaSerAlaGluTyTyIle 410
 |||||
 604 TATGTTATACGATGACAGATTTTAAAGCCAGCCGAGATACAAAT 653
 410 eLeuGluLysMetProGlnThrThrIleGlnValAspGlySerGluLysL 427
 |||||
 654 TCTCGAAGAAATGCTCAACACCATTCAGGTGATGGAGCGAGAAA 703
 427 yIleValSerIleLysAspPheLeuGlySerGluAsnLysArgThrgly 443
 |||||
 704 AATATGTCGATTAAGAGCTTCTTGAGAGTGAATATAAGAAACCGGG 753
 444 AlaLeuGlyAsnLeuLysAsnSerTySerTyAsnLysAspAsnGlu 460
 |||||
 754 GCGTTGGGTAATCTGAAAAACTCATCTCTTATATTAAGATTAATATGA 803
 460 uLeuSerHisPheAlaThrThrCysSerAspLysSerArgProLeuAsnA 477
 |||||
 804 ATTATCTGCTTGGCCACCACCTGCTCGGATTAAGTCAGCGCCGATCAAG 853
 477 spLeuValSerGlnLysThrThrglnLeuSerAspIleThSerArgPhe 493
 |||||
 854 ACTTGTTAGCCAAAAAACCACTGCTGCTGATTAATACACACAGTTT 903
 494 AsnSerAlaIleGluAlaLeuAsnArgPheIleGlnLysTyAspSerVa 510
 |||||
 904 AATTCAGCTATTAAGACCTGAACCGTTTCATTCGAAATATGATTCAGT 953
 510 lMetGlnArgLeuLeuAspAspThrSerGlyLys 521
 |||||
 954 GATGACAGCTCTGATGATGACACGCTGGTAA 987

seq_name: N_Geneseq_36:T38243

seq_documentation_block:

ID T38243 standard; DNA; 1014 BP.

AC T38243:

DT 28-DEC-1996 (first entry)

DE Y. pestis lcrv (V antigen) gene.

KW Plaque; vaccine; genetic immunisation; V antigen; lcrv;

KW F1 antigen; ds.

OS Yersinia pestis strain GB.

FH Key location/Qualifiers

FT cds 1..990

FT /*tag- a

FT misc.feature 1..10

FT /*tag- b

FT /*note- "vector pGEX-5x-2-derived bases"

FT mutation 16

FT /*tag- C

FT /*note- "base 16 is altered to a C from an

FT A to createan EcoRI site"

FT mutation 1006

FT /*tag- d

FT /*note- "base 1006 is altered to a T to create a

FT second in-frame stop codon"

FT mutation

FT W09628531-A1.

PN 19-SEP-1996.

PD 13-MAR-1996; G00571.

PE 13-MAR-1995; GB-005059.

PR 15-SEP-1995; GB-018946.

PR 05-DEC-1995; GB-024825.

PA (MTNA) UK SEC FOR DEFENCE.

PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
 DR WPI: 96-433824/43.

DR P-PSDB: W01041.
PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Claim 5; Page 32-35; 98pp; English.
CC A *lcrv* gene sequence (T38243) codes for the yersinia pestis V
CC antigen (W01041), which is capable of evoking protective immune
CC responses in animals. The gene was amplified from *Y. pestis*
CC DNA by PCR using primers (T33251 and T38239) homologous to the 5'
CC and 3' ends of the gene. The gene was inserted into vector
CC pSEX-5X-2, PMAL-p2 or PMAL-c2 (see also T38242) to allow prodn. of
CC recombinant V antigen for use in vaccines against plague.
CC Expression in gut-colonising organisms and attenuated *Salmonella*
CC typhi allows live vaccine prodn. F1/V antigen fusions were also
CC created (see also T38249 and T38256). The gene can itself be used
CC in genetic vaccines.
SQ Sequence 1014 BP; 343 A; 164 C; 205 G; 282 T;

```

alignment_scores:
  Quality: 1641.00      Length: 325
  Ratio: 5.049          Gaps: 0
Percent Similarity: 100.000  Percent Identity: 99.692

```

```
alignment_block:
```

US-08-6999-716A-2 X T38243

Align seg 1/1 to: T38243 from: 1 to: 1014

[illegible]

```

563  AATCCATGATGATAAATCCATTAATCTCATGATGATAAAATTTATATAGTTAT 61122
397  THRAGLUGLILIEPHELYSALSERIAGLUTRYLYSIIIELEUGULY 41333
613  ACAGATGAGAGATTTTTAAAGCCAGGCCAGACAGTACAAAATTCGCGAA 66202
413  smetProGlnThrThrIleGlnValAspGlySerGluLysLysIleValS 43030
663  AATGGCTCAAAACACACATTCAGTGGTGGATGGAGGACAAAATAATTACTCT 71122
430  erILeLYAspPheLeuGlySerGluAsnLysArGThrGlyAlaLeuGly 44630
713  CGATTAAGAGACTTCTCTTGAGTGAAGATGAGATAAAGAACCGGGCGCTGGGT 76230
447  AenLeuLysAsnSerTyrSerTyrAsnLysAspAsnAsnGluLeuSerH1 46333
763  AATCGAAAAAATCTACTACTCTTAATAATGAAGATATATATATATATATCTCA 81122
463  sPheAlaThrThrCysSerAspLysSerArgProLeuAsnAspLeuValS 48030
813  CTTTCCACACCACTGCTGCGATGAATGATCCAGGCCCTCAAGCACTGGTTA 86230
480  erGlnLysThrThrGlnLeuSerAspIleThrSerArgPheAsnSerAla 49630
863  GCCAAAAAACAACACTCAGCTGCTGTAATTAATCATCAGCTTTAATTCGCT 91122
497  IIEGLIALALEuAsnArgPheIIEGLInLysTyrAspSerValMetGlnAr 51333
913  AATGAGACACGAACCGTTTCAATTCACAGAAATATGATTCAGCTGATGCACGC 96230
513  gLeuLeuAspAspThrSerGlyLys 521
963  TCTGCTAGATGACACGCTGCTGTA 987
seq_name: N_Geneseq_36:104223

```

seq_name: N_Geneseq_36:T04223

seq_documentation_block:

ID	T04223	standard; DNA; 1014 BP.
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

AC T04223;
DT 19-APR-

DT 18-APR-1996 (first entry)
DE Partial Copy (Vantiden).c

DE: *palciat* LCRV (V antrigen) gene of Y. pestis.
KW: LCRV: V antigen: virulence: plaque: vaccine

OS
Yersinia pestis.
prague, vaccine, epirope; ss

Key

FT	cds	1. .990
----	-----	---------

ET

PN W09524475-A1.
PD 14-SEP-100E

PD 14-SEP-1995.
DE 06-MAR-1995. G00

PE 06-MAR-1995; GUL
PR 08-MAR-1994; GB-

PA (MINA) UK SEC E

PI Leary SEC, T1th

DR WPI; 95-328268/4

DR P-PSDB; R79962.

PT	Recombinant DNA
----	-----------------

PT Oral or parenteral
DS Clastm 6: Page 15

CC T04222-23 ARE DN

part of the Matu

expressed as a f

CC glutathione-S-t

CC is the highly v

CC animals, including

CC monomeric peptide

CC is postulated to
microorganisms

CC peptide are used

```

cc  repname are used
sq  sequence 1014

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

alignment_scores:
Quality: 1638.00

Ratio: 5.040 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.385

Alignment block:
US-08-699-716A-2 x T04223

Align seg 1/1 to: T04223 from: 1 to: 1014

```

197 Ileaargalatyrgluginasnproglinhispheilegialaspheugluly 213
13 ATTGAGGCTGAGAACAAACCAACATTTTGTGAGATCTAGAGAA 62
213 svaIaYValgluInleuthrghlyhISglYSerSeraVallenglugluL 230
63 AGTAGGGTGGAGACACTTACTGGTCATGCTTCTCATTAGAGAAAT 112
230 euValglInleuVallyspLysasnleuaspIleSeriIlelytYrSp 246
113 TGCTTCACTTGTCAAGATATAATATGATTTCCATTAAATATGAT 162
247 ProaYglYaspSerSergluValpheaIasnaryValIlethraspapiI 263
163 CCCAGAAAAGATTCGAGGTTTGGCCATAGAGAAATTAAGATAT 212
263 egluleuLeuLysIleleuValatyYrPheleuprogluaspThril 280
213 CGAATTCCTCAAGAAATCCTAGCTATTCTTACCCGAGATGCCATTC 262
280 euLysglYglYhIStyrAspasnghInleuasnghlyIlelytYrVal 296
263 TTAAGGGGCTATTATGACACCACTGCAAAATGGCATCAAGCGACTA 312
297 LysgluPheleuGluSerSeraSernthrgIntrpGluLeuArgAlaph 313
313 AAAAGATTCCTTGAATCATCCCGAATACACATGGCAATTCGGCGCTT 362
313 emetAlaValMetHisPheSerleuthrAlaaspArgIleaspasapi 330
363 CATGGCAGTATGATTCCTTAAACCGCGATCGTATCATGATGATGA 412
330 IeLeuValIleValaspSerMetasnHISglYAspAlaArgser 346
413 TTTGAAAGTATGATTGATTCATATGATCATGATGATGATGATGATG 462
347 LysleuArggluInleuValgluLeuValaspheleuThralasphe 363
463 AAGTGGCTGAGAAATTAAGTACAGTTCACCCCGAATTAAGATTTATTC 512
363 IValIleGlnAlaGluIleasnLysHISleuSerSeraSerglyThrIleA 380
513 AGTATTCACACCGAAATTAATAGCATCTGTAGTAGTGACACCATAA 562
380 snIleHisaspLysSerIleasnleuMetaspLysasnleuTygIlyYr 396
563 ATATCCAGAAATATCCATTAATCATGAGTAAATAATTTATATGATAT 612
397 ThraspGluGluIlePheLysAlaSerAlaGluTyYrIleleuGluTy 413
613 ACAGATGAGAGATTTTAAAGCCAGCCAGAGTACAAATATTCGAGAA 662
413 smetProglInthThrIleGlnValaspGlySeraIlyLysIleValS 430
663 AATGCTTCACCAACCATTCAGTGGATGGAGCGAGAAAATAATATGCT 712
430 erIleLysAspPheleuGlySeraLysasnLysArgThGlyAlaLeuGly 446
713 CGATTAAGAGACTTCTTGGAGTGAAGTAAAGAACCGGCGCTGGGT 762
447 AsnleuLysasnSerTyYrSeraLysasnLysaspasnghInleuSerH 463
763 AATCTGAAAACTCTACTCTTATATATAAGTAAATATGATTTATTCCTCA 812
463 sPhealathThrYsSeraspLysSeraArgProleuasnaspheuValS 480

```

```

|||||
813 CTTGCCCACACCGCTCGATAGTCCAGCCGCTCAACGACTTGTGTTA 862
480 ergInLysThrThrGlnLeuSeraspIleThrSerArgPheasnSera 496
863 GCCAAAACACTCAGCTGTGATATTAATACATGCTTTAATTCAGCT 912
497 IlegIuAlaLeuasnArgPheilegInLysTyYrAspSeraValmetGlnar 513
913 ATGAACACTGAACCGTTTCATTGAGAAATATGATTCATGATGACAG 962
513 gLeuLeuaspSerpThrSeraLys 521
963 TCTCTAGATGACACGCTGTGTA 987

```

seq_name: N_Geneseq_36:Q92819

```

seq_documentation_block:
ID Q92819 standard; DNA; 542 BP.
AC Q92819;
DR 17-DEC-1995 (first entry)
DE Yersinia pestis cafI (FI) antigen in plasmid pPoreFb.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
KM bubonic plague; pneumonic plague; ds.
OS Yersinia pestis.
FH key
FT Location/Qualifiers
   2..7
   /tag= a
   /note= "first protein encoded by pPoreFb"
FT
FT misc_feature
   1..6
   /tag= b
FT
FT misc_feature
   536..541
   /tag= c
FT
FT /note= "cafI open reading frame downstream seq."
FT
FT cds
   21..530
   /tag= d
   /note= "cafI fusion protein"
FT
FT WO9518231-A1.
PD 06-JUL-1995.
PN 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PI (MINA) UK SEC FOR DEFENCE.
PA Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 95-246396/32.
DR P-PSDB: R76528.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response.
PT against Yersinia pestis, at mucosal surfaces.
PS Disclosure; Page 19-20; 27pp; English.
CC The sequence represents the plasmid pPoreFb including the entire
CC Y. pestis cafI (FI) antigen gene having a 5' tail including a SacI
CC restriction site, and up to 1000 downstream of the cafI ORF.
CC The DNA construct can be used to transform human or animal gut
CC colonizing microorganisms, specifically attenuated Salmonella
CC typhimurium or Salmonella typhi. The transformed microorganisms
CC can be used as live/attenuated vaccines which induce immune
CC responses at mucosal surfaces. The vaccines provide protection
CC against infection with Y. pestis, and are parenterally and orally
CC active vaccines offering protection against bubonic and pneumonic
CC plague.
SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

```

alignment_scores:

Quality: 852.00 Length: 171
Ratio: 4.982 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.415

alignment_block:
US-08-699-716A-2 x Q92819

Align seg 1/1 to: Q92819 from: 1 to: 542

23 HisMetLysLysLysSerSerValIleAlaIleAlaLeuPheGlyThrI1 39
 18 TATATGAAAAAATCATGCTCCGTATCGCCATGCAATTATTTGGAACTAT 67
 39 eAlaThrAlaAsnAlaAlaAspLeuThrAlaSerThrThrAlaThrAlaT 56
 68 TCCAACTGCTAATGCGGAGATTTACTGCACACCACTGCACAGCGCA 117
 56 hLeuValGluProAlaArgIleThrLeuThrTyrLysGluGlyAlaPro 72
 118 CTCTGTTGAAACGAGCCGCACTACCTTACATATATAGAAAGCGCTTCA 167
 73 IleThrIleMetAspAsnGlyAsnIleAspThrGluLeuValGlyTh 89
 168 ATTACAAATTATGACATGAAACATGATACAGATTAATCTTGTGTAC 217
 89 rLeuThrLeuGlyLysLysThrGlyThrThrSerThrSerValAsp 106
 218 GCTTACTCTGGCGGCTATAAACAGAAACCACTACACATCTGTTAACT 267
 106 hEThrAspAlaAlaGlyAspPrometTyrLeuThrPheThrSerGlnAsp 122
 268 TTACAGATGCCGCGGTATCCATCTACTTAACATTTACTTCTCAGAT 317
 123 GlyAsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArgAs 139
 318 GGAATTAACACCAATCTCACTCAAAAGATGGCAAGATTCTAGAGA 367
 139 pPheAspIleSerProLysValAsnGlyLysLeuValGlyAspAsp 156
 368 TTTGATATCTCTCCTAAGTAAACGATGAGAACTTGIGGGGATGAG 417
 156 aValLeuAlaThrGlySerGlnAspPhePheValArgSerIleGlySer 172
 418 TCGTCTTGCTAGCGGAGCCAGAGATTCTTGTGCTCAATGGTCC 467
 173 LysGlyGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 189
 468 AAAGCGGTAAACTGACAGAGTAATCACTGATGCTGTAACCTGTAAC 517
 189 rValSerAsnGln 193
 518 CGTATCTAACCA 530
 seq_name: N_Geneseq_36:V41594
 seq_documentation_block:
 ID V41594 standard; DNA; 544 BP.
 AC V41594;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nyf1(a)sec544.
 KW F1 antigen; plasmid; vaccine; plaque; ds.
 OS Yersinia pestis.
 FH Key Location/Qualifiers
 FT CDS 17..529
 FT /*tag= a
 FT /product= "F1 antigen"
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EV, Osorio JE, Thomas RE;
 DR WPI; 98-33331/29.
 PT Protection of animals against plague - using nucleic acid encoding
 PS antigen from Yersinia, Pasteurella and Francisella spp.
 CC Claim 8; Page 51; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.

SO Sequence 544 BP; 164 A; 120 C; 114 G; 146 T;
 alignment_scores:
 Quality: 852.00 Length: 171
 Ratio: 4.982 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.415
 alignment_block:
 us-08-699-716a-2 x V41594
 Align seg 1/1 to: V41594 from: 1 to: 544
 23 HisMetLysLysLysSerSerValIleAlaIleAlaLeuPheGlyThrI1 39
 18 TATATGAAAAAATCATGCTCCGTATCGCCATGCAATTATTTGGAACTAT 67
 39 eAlaThrAlaAsnAlaAlaAspLeuThrAlaSerThrThrAlaThrAlaT 56
 68 TCCAACTGCTAATGCGGAGATTTACTGCACACCACTGCACAGCGCTCA 113
 56 hLeuValGluProAlaArgIleThrLeuThrTyrLysGluGlyAlaPro 72
 114 CTCTGTTGAAACGAGCCGCACTACCTTACATATATAGAAAGCGCTTCA 163
 73 IleThrIleMetAspAsnGlyAsnIleAspThrGluLeuValGlyTh 89
 164 ATTACAAATTATGACATGAAACATGATACAGATTAATCTTGTGTAC 213
 89 rLeuThrLeuGlyLysLysThrGlyThrThrSerThrSerValAsp 106
 214 GCTTACTCTGGCGGCTATAAACAGAAACCACTACACATCTGTTAACT 263
 106 hEThrAspAlaAlaGlyAspPrometTyrLeuThrPheThrSerGlnAsp 122
 264 TTACAGATGCCGCGGTATCCATCTACTTAACATTTACTTCTCAGAT 313
 123 GlyAsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArgAs 139
 314 GGAATTAACACCAATCTCACTCAAAAGATGGCAAGATTCTAGAGA 363
 139 pPheAspIleSerProLysValAsnGlyLysLeuValGlyAspAsp 156
 364 TTTGATATCTCTCCTAAGTAAACGATGAGAACTTGIGGGGATGAG 413
 156 aValLeuAlaThrGlySerGlnAspPhePheValArgSerIleGlySer 172
 414 TCGTCTTGCTAGCGGAGCCAGAGATTCTTGTGCTCAATGGTCC 463
 173 LysGlyGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 189
 464 AAAGCGGTAAACTGACAGAGTAATCACTGATGCTGTAACCGTAAAC 513
 189 rValSerAsnGln 193
 514 CGTATCTAACCA 526
 seq_name: N_Geneseq_36:V41596
 seq_documentation_block:
 ID V41596 standard; DNA; 544 BP.
 AC V41596;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nyf1sec510.
 KW F1 antigen; plasmid; vaccine; plaque; ds.
 OS Yersinia pestis.
 FH Key Location/Qualifiers
 FT CDS 17..532
 FT /*tag= a
 FT /product= "F1 antigen"
 PN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.

PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haenes EJ, Osorio JE, Thomas RE;
 DR WPI; 98-333331/29.
 DR P-PSDB; W59783.
 PT Protection of animals against plague - using nucleic acid encoding
 antigen from *Yersinia*, *Pasteurella* and *Francisella* spp.
 PS Claim 8; Pages 33-34; 75pp; English.
 CC This is the nucleotide sequence of a *Yersinia* pestis F1 antigen, used
 in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially *Yersinia* pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 protecting an animal from contracting plague.
 SQ Sequence 344 bp; 166 A; 118 C; 112 G; 148 T;

alignment_scores:
 Quality: 852.00 Length: 171
 Ratio: 4.982 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.415

alignment_block:
 US-08-699-716a-2 x V41596 ..

Align seg 1/1 to: V41596 from: 1 to: 544

```

23 Hismetlyslsileserservaliiealalealeupheglythrll 39
14 TATATGAAAAATCAGTCCGTTATCCGATTCATTTATTTGAACTAT 63
39 ealathralaasnalaalaspheuthralaserrhrthralathralat 56
64 TGCACATGCAATGCGGACATTTAATGCAACACACATGCAACGCA 113
56 hrlenuvalguproalaargillethleuthrlyrlysglugialapro 72
114 CTCTGTGTGACACGCCGATCCTTACATATTAAGAGGCGCTCCA 163
73 lilethrillemetaspasnnglyasniileaspthrliueneuvalglyth 89
164 ATTACATTTTGACATGGAACATCGATACACAAATTACTTGTGTGATC 213
89 rleuthrleuglyglytyrlysthrlythrlythrlythrlythrlythr 106
214 GCTTACTCTGGCGGCTATAAAGACAGACACATGCTGTTACT 263
106 hehthraspalaalaglyaspProwmetrlyrleuthrphetrserglaasp 122
264 TTACAGATGCGGCGGTGATCCCATGCTTACTTACTTCTCAGAGAT 313
123 G1YAsnashisglnphetrhrthlyrlyvalilleglylyaspserrargas 139
314 GGAATATACCAACCAATTCATCAAAAGATGCGAAGATTCCTTAAGA 363
139 pheaspilseerProlyvalasnnglyluasnleuvalglyaspaspy 156
364 TTTTGAATATCTCTCCTAAGGTAAACGCTGAAACCTTGTGGGATACG 413
156 alvalleualathnglyserglnasphethevalargserilleglyser 172
414 TCGTCTTGCGTACGGGACGCCAGGATTTCTTGTCTCCATATGTGTCC 463
173 lygilyglylystleuualaglylystlythrthraspalaalathrvalth 189
464 AAAGGCGGTAACTGACGACAGTAATACATGATGCTGTAAACCGTAAC 513
189 rValaserasnngln 193
514 CGTATCTAACCA 526
seq_name: N_Geneseq_36:T38248
seq_documentation_block:

```

ID T38248 standard; DNA; 547 bp.
 AC T38248;
 DE 28-DEC-1996 (first entry)
 DR Y. pestis F1 antigen caf1 gene (including signal sequence).
 KW Plague vaccine; genetic immunisation; F1 antigen; caf1;
 V antigen; ds.
 OS *Yersinia* pestis strain GB.
 FH Key Location/Qualifiers
 FT signal_peptide 29..91
 FT mat_peptide 92..538
 FT /*tag= a
 FT /*tag= b
 FM W09628551-A1.
 PD 19-SEP-1996.
 PR 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MVA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
 DR WPI; 96-433824/43.
 DR P-PSDB; W01043.
 PT *Yersinia* pestis V antigen and F1 antigen or their protective
 epitopic parts - useful in vaccine for protection against plague
 PS Claim 41; Page 61-62; 98pp; English.
 CC A DNA sequence (T38248) comprises the caf1 gene, including the
 CC signal sequence, coding for the F1 antigen (W01043) of *Yersinia*
 CC pestis. It was obt. by PCR amplification (see also T38257-58)
 CC of Y. pestis DNA. The PCR product was cloned into plasmid pBKCMV
 CC and the resulting plasmid (pF1AB) was used to transform E. coli Nova
 CC Blue. Purified plasmid, when administered by i.m. injection,
 CC induced an immunoglobulin response to F1 in Balb/c mice. Live
 CC vaccines comprising gut colonising organisms transformed with the
 CC caf1 gene (see also T38244) can be used to protect a host animal
 CC against plague.
 SQ Sequence 347 bp; 165 A; 120 C; 115 G; 147 T;

alignment_scores:
 Quality: 852.00 Length: 171
 Ratio: 4.982 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.415

alignment_block:
 US-08-699-716a-2 x T38248 ..

Align seg 1/1 to: T38248 from: 1 to: 547

```

23 Hismetlyslsileserservaliiealalealeupheglythrll 39
26 TATATGAAAAATCAGTCCGTTATCCGATTCATTTATTTGAACTAT 75
39 ealathralaasnalaalaspheuthralaserrhrthralathralat 56
76 TGCACATGCAATGCGGACATTTAATGCAACACACATGCAACGCA 125
56 hrlenuvalguproalaargillethleuthrlyrlysglugialapro 72
126 CTCTGTGTGACACGCCGATCCTTACATATTAAGAGGCGCTCCA 175
73 lilethrillemetaspasnnglyasniileaspthrliueneuvalglyth 89
176 ATTACATTTTGACATGGAACATCGATACAGAAATTACTTGTGTGATC 225
226 GCTTACTCTGGCGGCTATAAAGACAGACACATGCTGTTACT 275
89 rleuthrleuglyglytyrlysthrlythrlythrlythrlythrlythr 106
106 hehthraspalaalaglyaspProwmetrlyrleuthrphetrserglaasp 122
276 TTACAGATGCGGCGGTGATCCCATGCTTACTTACTTCTCAGAGAT 325
123 G1YAsnashisglnphetrhrthlyrlyvalilleglylyaspserrargas 139

```

```

|||||
326 GGAATATACCAACCAATTCATACAAAGTGAATGGCAGAGATTCAGAGA 375
139 pPhaspiIleSerProlyValasnGlyluasnLeuValGlyaspav 156
376 TTTTGATATCTCTCTTAAGGTAAACGGTAGAACCTTGTGGGGGATGACG 425
156 aValleuAlaThrGlySerGlnaspPhepheValArgSerIleGlySer 172
426 TCGTCTTGCTAGCGGACGACGAGATTTCTTGTCTGCTCAATGTGCTCC 475
173 LysGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 189
476 AAAGCGGTAAACTTCGACGACGAGTAATAATACACTGATGCTGTAAACGTAAC 525
189 rValSerAsnGln 193
526 CGTATCTAACCAA 538

```

seq_name: N_Geneseq_36.V41595

seq_documentation_block:

```

ID V41595 standard; DNA; 510 BP.
AC V41595;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of FI antigen nYp1(b)sec544.
KW FI antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
PN WO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
PT Protection of animals against plague - using nucleic acid encoding
PS antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 53; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis FI antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 510 BP; 153 A; 114 C; 106 G; 137 T;

```

alignment_scores:

```

Quality: 850.00 Length: 170
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-08-699-716A-2 x V41595 ..

Align seg 1/1 to: V41595 from: 1 to: 510

```

24 MetLysLysIleSerSerValIleAlaIleAlaLeuPheGlyThrIleAl 40
|||||
1 ATGAAAAAATCACTCCGTTATCGCATTCATATTTGGAACATATGC 50
40 aThrAlaAsnAlaAlaAspLeuThrAlaSerThrThrAlaThrAlaThrL 57
|||||
51 AACGCTAAATCGGCGAGATTTAACTGCAAGCACCACTGCAACGCGCACTC 100
57 euValGluProAlaArgIleThrLeuThrTyrLysGluGlyAlaProIle 73
|||||
101 TTGTTGAACACGCGCATCTCTTACATATAAGAAAGCGCTCCATTT 150
74 ThrIleLeuAspAsnGlyAsnIleAspThrGluLeuValGlyThrLe 90
|||||
151 ACAATATATGACATGGAACATCATACAGATAATATTACTTGTGGAGCT 200
90 uThrLeuGlyGlyTyrLysThrGlyThrThrSerThrSerValAsnPheT 107

```

```

|||||
201 TACTCTTGCGGCTATTAACAGAACACCTAGCACATCTTTACTTTA 250
107 hrAspAlaAlaGlyAspPrometTyrLeuThrPheThrSerGlnaspGly 123
251 CAGATGCGCGGCGTGATGCCATGTACTTAACATTTACTCTCAGAGAGA 300
124 AsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArgAspPh 140
301 AATAACCAACCAATTCATCAAAAGTGAATTGCAAGGATTTCTAAGATTT 350
140 eAspIleSerProlyValasnGlyluasnLeuValGlyAspAspVal 157
351 TGATATCTCTCTTAAGGTAAACGGTAGAACCTTGTGGGGATGACGTCG 400
157 aLeuAlaThrGlySerGlnaspPhepheValArgSerIleGlySerLys 173
401 TCTTGCTAGCGGACGACGAGATTTCTTGTCTGCTCAATGTGTCANA 450
174 GlyGlyLysLeuAlaAlaGlyLysTyrThrAspAlaValThrValTh 190
451 GCGGTAAACTTCGACGACGAGTAATAATACACTGATGCTTAACCGTACGT 500
190 lSerAsnGln 193
501 ATCTAACCAA 510

```

seq_name: N_Geneseq_36.V41609

seq_documentation_block:

```

ID V41609 standard; DNA; 447 BP.
AC V41609;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of FI antigen nYp1mat447.
KW FI antigen; plasmid; vaccine; plaque; ds.
OS Yersinia pestis.
FH key Location/Qualifiers
FT CDS
FT FT 1..447
FT /tag= a
FT /product= "FI antigen"
PN WO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR P-PsDB; W59788.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 64; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis FI antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 447 BP; 133 A; 102 C; 95 G; 117 T;

```

alignment_scores:

```

Quality: 755.00 Length: 149
Ratio: 5.067 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-08-699-716A-2 x V41609 ..

Align seg 1/1 to: V41609 from: 1 to: 447

```

45 AlaAspLeuThrAlaSerThrThrAlaThrAlaThrLeuValGluProAl 61
|||||
1 GCAGATTTAACTGCAAGCACCACTGCAACGCAACTCTTGTGAACAGC 50

```

```

61 aargllethleuthrtyrlysgllyalaproillethrllemetaspa 78
|||||
51 CGGCATCACTCTTACATATAAGGAAGCGCTCCAAATTACATATATGACA 100
|||||
78 snglyasnlleasphrgluleuleuvalglythrlleuthrleuglygly 94
|||||
101 AVGGAAACATCATACAGAAATTACTGTGTGACCTTCTTGCGCGC 150
|||||
95 tyrlvsthrlythrtserthrservalasphethraspalaaagl 111
|||||
151 TATATAAACGGAACACATGACATCTGTAACTTACAGATGCCGCGG 200
|||||
111 yasprometlyrleuthrphethrserclinaspglyasnaenhsiglup 128
|||||
201 TGATCCCATGACTTACATTAATTACTTCTCAGATGGAATAACCAACCAT 250
|||||
128 hehthrlthrlsvallellyllyaspserserargaspheasplleserpro 144
|||||
251 TCACATCAAAAGTAGATTGGCAAGATTCTAGAGATTTGTATATCTCTCT 300
|||||
145 lysvalasnglyglusnleuvalglyaspaspvalvalleuathrcgl 161
|||||
301 AAGGTAAAGGTCAGAACCTTGTGGGGATGACGCTGCTGGCTACGGG 350
|||||
161 yserglnasphphevalargserilleglyserlysglyglylysleua 178
|||||
351 CAGCCAGATTTCTTGTTCGCCCAATGGTTCCAAAGCGGTAACCTTG 400
|||||
178 laalaglystfyrthraspalavalthvalthvalserasngln 193
|||||
401 CAGCAGTAATACACTGATGCTGTACCGTAACGTAATCTAACCAA 447
|||||

```

seq_name: N_Geneseq_36:V41601

seq_documentation_block:

ID V41601 standard; DNA; 450 BP.

AC V41601;

DT 26-OCT-1998 (first entry)

DE Nucleotide sequence of F1 antigen nyFlnat450.

KW F1 antigen; plasmid; vaccine; plaque; ds.

OS Yersinia pestis.

PN MO9824912-A2.

PD 11-JUN-1998.

PF 04-DEC-1997; U22617.

PR 04-DEC-1996; US-767115.

PA (HESK-) HESKA CORP.

PI Haanes EJ, Osorio JE, Thomas RE;

DR WPI; 98-333331/29.

PT Protection of animals against plague - using nucleic acid encoding

pr antigen from Yersinia, Pasteurella and Francisella spp.

PS Claim 8; Page 61; 75pp; English.

CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used

CC in the method of the invention. Plasmid and host cells are used to

CC produce recombinant antigens, especially Yersinia pestis antigens.

CC The recombinant antigens can be used in vaccines that are capable of

CC protecting an animal from contracting plague.

Sequence 450 BP; 134 A; 102 C; 96 G; 118 T;

alignment_scores:

Quality: 755.00

Ratio: 5.067

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-08-699-716A-2 x V41601

Align seg 1/1 to: V41601 from: 1 to: 450

45 AlaaspleuthrthraserthrtthralatthAlatthrlleuvalgluproal 61
 |||||||
 4 GCAGATTTAATGCAAGCACCACTGCAAGCGCAACTCTGTGTGAAACGAG 53

```

61 aargllethleuthrtyrlysgllyalaproillethrllemetaspa 78
|||||
54 CGGCATCACTCTTACATATAAGGAAGCGCTCCAAATTACATATATGACA 103
|||||
78 snglyasnlleasphrgluleuleuvalglythrlleuthrleuglygly 94
|||||
104 AVGGAAACATCATACAGAAATTACTGTGTGACCTTCTTGCGCGC 153
|||||
95 tyrlvsthrlythrtserthrservalasphethraspalaaagl 111
|||||
154 TATATAAACGGAACACATGACATCTGTAACTTACAGATGCCGCGG 203
|||||
111 yasprometlyrleuthrphethrserclinaspglyasnaenhsiglup 128
|||||
204 TGATCCCATGACTTACATTAATTACTTCTCAGATGGAATAACCAACCAT 253
|||||
128 hehthrlthrlsvallellyllyaspserserargaspheasplleserpro 144
|||||
254 TCACATCAAAAGTAGATTGGCAAGATTCTAGAGATTTGTATATCTCTCT 303
|||||
145 lysvalasnglyglusnleuvalglyaspaspvalvalleuathrcgl 161
|||||
304 AAGGTAAAGGTCAGAACCTTGTGGGGATGACGCTGCTGGCTACGGG 353
|||||
161 yserglnasphphevalargserilleglyserlysglyglylysleua 178
|||||
354 CAGCCAGATTTCTTGTTCGCCCAATGGTTCCAAAGCGGTAACCTTG 403
|||||
178 laalaglystfyrthraspalavalthvalthvalserasngln 193
|||||
404 CAGCAGTAATACACTGATGCTGTACCGTAACGTAATCTAACCAA 450
|||||

```

seq_name: N_Geneseq_36:V41600

seq_documentation_block:

ID V41600 standard; DNA; 474 BP.

AC V41600;

DT 26-OCT-1998 (first entry)

DE Nucleotide sequence of F1 antigen nyFlnat474.

KW F1 antigen; plasmid; vaccine; plaque; ds.

OS Yersinia pestis.

PN Location/Qualifiers

FM key

FT CDS

FT /*tag= a

FT /product= "F1 antigen"

PN MO9824912-A2.

PD 11-JUN-1998.

PF 04-DEC-1997; U22617.

PR 04-DEC-1996; US-767115.

PA (HESK-) HESKA CORP.

PI Haanes EJ, Osorio JE, Thomas RE;

DR WPI; 98-333331/29.

DR P-PSDB; W59787.

PT Protection of animals against plague - using nucleic acid encoding

pr antigen from Yersinia, Pasteurella and Francisella spp.

PS Claim 8; Pages 59-60; 75pp; English. Yersinia pestis F1 antigen, used

CC in the method of the invention. Plasmid and host cells are used to

CC produce recombinant antigens, especially Yersinia pestis antigens.

CC The recombinant antigens can be used in vaccines that are capable of

CC protecting an animal from contracting plague.

Sequence 474 BP; 143 A; 106 C; 99 G; 126 T;

alignment_scores:

Quality: 755.00

Ratio: 5.067

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-08-699-716A-2 x V41600

Align seg 1/1 to: V41600 from: 1 to: 474

```

45 AlaSpleuThrAlaSerThrThrAlaThrAlaThrLeuValGluProAl 61
   |||||||
10 GCACATTACTGCAAGCACCACGCAACGGCAACTCTGTTGAACCGC 59
   |||||||
61 ArgIleThrLeuThrTyrTyrGluGlyAlaProIleThrIleMetAsp 78
   |||||||
60 CCGCATCAGCTTACATATAGGAAGCGCCCATTCATTCATTGGACA 109
   |||||||
78 snGlyAsnIleAspThrGluLeuLeuValGlyThrLeuThrLeuGly 94
   |||||||
110 ATGGAACATCGATACAGATTACTGTTGGTAGCTTACTCTGGCGGC 159
   |||||||
95 TyrIleThrGlyThrThrSerThrSerValAsnThrAspAlaAlaG 111
   |||||||
160 TATTAACAGGACCACTAGCAGCATCTGTTACTTACAGATGCGCGGG 209
   |||||||
111 YAspPrometYrLeuThrPheThrSerGlnAspGlyAsnAsnHsGlnP 128
   |||||||
210 TGATCCATGACTTACATTACTTCTCAGATGGAATACCAACCAAT 259
   |||||||
128 heThrThrLysValIleGlyLysAspSerArgAspPheAspIleSerPro 144
   |||||||
260 TCACTACAAAGTAGTAGCAAGATCTAGAGATTTGATATCTCTCCCT 309
   |||||||
145 LysValAsnGlyGluAsnLeuValGlyAspAspValValLeuAlaThrG 161
   |||||||
310 AAGGTAAACGGTGAGAACCTTGTGGGGATGACCTCGCTTGCTACGGG 359
   |||||||
161 YSerGlnAspPhePheValArgSerIleGlySerLysGlyGlyLysLeu 178
   |||||||
360 CAGCGAGATTCTTGTTCCTCATTTGTTCCAAAGCGCGTAACCTG 409
   |||||||
178 laaIaGlyLysTyrThrAspAlaValThrValThrValSerAsnGln 193
   |||||||
410 CAGCGGTAAATACATGATGCTTAACCGTAACCGTATCAACCAA 456

```

seq_name: N_Geneseq_36:Q92817

seq_documentation_block:

```

ID 092817 standard; DNA; 541 BP.
AC 092817;
DE 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen in plasmid pFGAL2.
KM Vaccine; antigen; Salmonella typhimurium; Salmonella typh;
KM bubonic plague; pneumonic plague; ds.
OS Yersinia pestis.
FH Key
   Location/Qualifiers
FT cds
   2..454
   /*tag= a
FT misc_feature
   1..6
   /*tag= b
FT misc_feature
   536..541
   /note= "lacZ promoter fusion site"
FT FT
   /*tag= c
   /note= "vector pFGAL2 bases"
PN MO9518231-A1.
PD 06-JUL-1995.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA) UK SEC FOR DEFENCE.
PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
PI WPI: 95-246396/32.
DR P-ESDB: R76526.

```

PT DNA constructs capable of transforming microorganisms - which can be used as live or attenuated vaccines which induce an immune response, against Yersinia pestis, at mucosal surfaces.

PS Claim 7: Page 15-16; 27pp; English.

CC The sequence represents the plasmid pFGAL2 construct showing the fusion of the first few bases of beta-galactosidase in the vector with the Y. pestis cafi (F1) antigen minus its signal sequence and having a 5' tail including a SacI restriction site, and up to the cafi AACG-3' end with some vector bases. The DNA construct can be

CC used to transform human or animal gut colonizing microorganisms, specifically attenuated Salmonella typhimurium or Salmonella typh. CC The transformed microorganisms can be used as live/attenuated CC vaccines which induce immune responses at mucosal surfaces. The CC vaccines provide protection against infection with Y. pestis, and CC are parenterally and orally active vaccines offering protection CC against bubonic and pneumonic plague.

Sequence 541 bp; 163 A; 120 C; 111 G; 147 T;

alignment_scores:

Quality: 755.00 Length: 151

Percent Similarity: 100.000 Percent Identity: 98.013

alignment_block:

US-08-699-716A-2 x Q92817

Align seg 1/1 to: Q92817 from: 1 to: 541

```

43 AsnAlaAlaSpleuThrAlaSerThrThrAlaThrAlaThrLeuValG 59
   |||||||
2 AGCTGGCAGATTAACTGCAAGCACCACGCAACGGCAACTCTGTTGA 51
   |||||||
59 UProlAlaArgIleThrLeuThrTyrTyrGluGlyAlaProIleThrIle 76
   |||||||
52 ACCAGCCCGGATCATCTTACATTAAGGAAGCGCCCTCAATTAACAAT 101
   |||||||
76 eAspAsnGlyAsnIleAspThrGluLeuLeuValGlyThrLeuThrLeu 92
   |||||||
102 TGGCAATGCAACATCGATACAGATTACTGTTGGTAGCTTACTCTT 151
   |||||||
93 GlyIleTyrIleThrGlyThrThrSerThrSerValAsnThrAspAl 109
   |||||||
152 GCGGCTATAAAGCAGACACGACATCTGTTAACTTACAGATGC 201
   |||||||
109 aAlaGlyAspPrometYrLeuThrPheThrSerGlnAspGlyAsnAsn 126
   |||||||
202 CGCGGTGATCCCATGACTTACATTACTTCTCAGATGGAATTAAC 251
   |||||||
126 lAsnIlePheThrLysValIleGlyLysAspSerArgAspPheAspIle 142
   |||||||
252 ACCAATTCATCAAAAGTAGTAGCAAGATCTTACAGATTTGATATC 301
   |||||||
143 SerProLysValAsnGlyGluAsnLeuValGlyAspAspValValLeu 159
   |||||||
302 TCTCTTAAGGTAAACGGTGAGAACCTTGTGGGGATGACGCTTGGC 351
   |||||||
159 aThrGlySerGlnAspPhePheValArgSerIleGlySerLysGlyL 176
   |||||||
352 TACGGGAGCGCAGGATTTCTTGTCCCTCAATTGGTTCCAAAGCGGTA 401
   |||||||
176 YLeuAlaAlaGlyLysTyrThrAspAlaValThrValThrValSerAs 193
   |||||||
402 AACTTGCAGCAGTAATACATGATGCTTAACCGTAACCGTATCTAAC 451
   |||||||
193 Gln 193
   |||
452 CAA 454

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Percent Similarity: 99.004 Percent Identity: 98.606

alignment_block:

US-08-699-716a-2 x US-08-913-477-22

Align seg 1/1 to: US-08-913-477-22 from: 1 to: 1530

```

23 HisMetLysLysLieserSerValIleAlaIleAlaLeuPheGlyThrI1 39
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
10 TATATGAAAAAATCAGTCCGTTATCCCATTCGCTATTATTGGAACTAT 59
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
39 eAlaThrAlaAsnAlaAlaAspLeuThrAlaSerThrAlaThrAlaT 56
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
60 TGCACACTGCTATGCGGAGATTTAAGTCAAGACCCACTGCAAGCGAA 109
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
56 hLeuValGluProAlaArgIleThrLeuThrIleGlyGluAlaPro 72
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
110 CTCTGTGTGAACGACCGCATCTTACATATTAAGGAAGCGCTCCA 159
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
73 IleThrIleMetAspAsnGlyAsnIleAspThrGluLeuValGlyTh 89
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
160 ATTACAAATTATGAGCAATGAAACATCATACAGATTTCTGTGTGATC 209
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
89 rLeuThrLeuGlyGlyIleYrLysThrGlyThrSerThrSerValAsp 106
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
210 GCTTACTCTTGGCGGCTATTAACAGGACCACTAGCACATCTGTAACT 259
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
106 hethrAspAlaAlaGlyAspProMetIleLeuThrPheThrSerGln 122
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
260 TTACAGATGCGCGGTATCCCATGTTACTTACATTTACTTCAGAGAT 309
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
123 GlyAsnAsnHisGlnPheThrThrLysValIleGlyLysAspSerArg 139
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
310 GGAATATACCCACCAATCTACTACAAAGATGTGGCAAGATTTCTAGA 359
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
139 pPheAspIleSerProLysValAsnGlyLysLeuValGlyAspAspV 156
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
360 TTTTGATATCTCTCTTAAGGTAAACGGTGAACCTTGGGGGATGACG 409
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
156 aValIleuAlaThrGlySerGlnAspPheValArgSerIleGlySer 172
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
410 TCGTCTGGCTACGGGACCGAGATTTCTTGTGCTCATTTGTTGCC 459
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
173 LysGlyGlyLysLeuAlaAlaGlyLysIleThrAspAlaValIleTh 189
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
460 AAAGCGGTAACTGACGAGGTAATATCATGCTGTACCGGTATAC 509
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
189 rValSerAsnGln.....GluPheMetIleArgAlaIleArgGluGln 203
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
510 CGTATCTTACCAAGATCCATCGAAGGTCGTATTAGACCTTAGCAACA 559
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
203 snProGlnHisPheIleGluAspLeuGlyLysValArgValGluGln 219
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
560 ACCCAACAACATTTATTGAGATCTAGAAAAGTAGGGTGAACAACCT 609
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
220 ThrGlyHisGlySerSerValLeuGluGluLeuValGlnLeuValLys 236
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
610 ACTGCTCATGTGCTCTCATGTTTAGAAGAAATGTTGCTAGTAGCAAA 659
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
236 pLysAsnIleAspIleSerIleLysTyrAspProArgLysAspSerGly 253
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
660 TAAAAATATATATATTTCCATTAATATATGATCCCAAGAAAATTCGG 709
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
253 aPheAlaAsnArgValIleThrAspAspIleGluLeuLeuLysLysIle 269
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
710 TTTTTCGAATAGAGTATTAAGTATGATATCGAATTCCTCAAGAAATC 759
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
270 LeuAlaIleThrLeuProGluAspThrIleLeuLysGlyLysIleThr 286
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
760 CTACTATTATTCTACCCGAGAGTCCATCTTAAAGCGGTCTATTAGA 809
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
286 pAsnGluLeuGlnAsnGlyIleLysArgValLysGluPheLeuGlnSer 303
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

```

810 CAACCACTGCAAAATGCAATCAAGCGATGAAGAGTCTCTGATCAT 859
303 erProAsnThrGlnIlePgluLeuArgAlaPheMetAlaValMetHisPhe 319
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
860 CGCCGAATACACAAATGGGAATTTGGGGCGTTTCATGCGAGATTAATCAT 909
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
320 SerLeuThrAlaAspArgIleAspAspAspIleLeuLysValIleValAs 336
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
910 TCTTTAACCCCGGATCGTATCATGATGATGATATTTTGAAGATGATGTTGA 959
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
336 pSerMetAsnHisGlyAspAlaAspSerLysLeuArgGluGluLeuAs 353
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
960 TTCAATGATCATCATGCTGATGCTGATGCCCTACCACTTCGTGAAGAAAT 1009
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
353 lAgIleuThrAlaGluLeuLysIleTyrSerValIleGlnAlaGluIle 369
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1010 CTGAGCTTACCGCCGAATTAAGATTTATTCAGTTATTCAAGCGGAAT 1059
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
370 AsnLysHisLeuSerSerSerGlyThrIleAsnIleHisAspLysSer 386
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1060 AATAAGCATCTGTCTAGTAGTGCACCATTAATATCCATATTAATCCAT 1109
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
386 eAsnLeuMetAspLysAsnLeuIleYrGlyIleThrAspGluGluIlePhe 403
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1110 TAACTCATGATTAATAATTTATTTATGTTATACAGATGAAGAGATTTT 1159
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
403 yAlaSerAlaGluIleYrLysIleLeuGluLysMetProGlnThrIle 419
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1160 AAGCCAGCGAGATACAAATTCGAGAAATTCGAAAGCTCAACACCAT 1209
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
420 GlnValAspGlySerGlyLysIleValSerIleLysAspPheLeuGlu 436
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1210 CAGGTGATGGAGCGAGAAAAAATGCTCGATAAGACATTTCTTTGG 1259
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
436 ySerGluAsnLysArgThrGlyAlaLeuGlyAsnLeuLysAsnSerTyr 453
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1260 AAGTGAGATTAAGAAACCGGGCGGTGGGTAACTCTAAAAACTCATACT 1309
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
453 erTyrAsnLysAspAsnAsnGluLeuSerHisPheAlaThrIleCysSer 469
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1310 CTATTAATTAAGATTAATGAATATATCTCATCTTGCACACACTGCTCG 1359
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
470 AspLysSerArgProLeuAsnAspLeuValSerGlnLysThrIleGln 486
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1360 GATAAGTCCAGCGCGCTCAACGACTGTGTAAGCCAAAAACAACCTGACT 1409
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
486 uSerAspIleThrSerArgPheAsnSerAlaIleGluAlaLeuAsnArg 503
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1410 GTCTGATATTCATCATACGTTTAAATTAAGCTATTGAAGCACTGAACCG 1459
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
503 heIleGlnLysTyrAspSerValMetGlnArgLeuLeuAspAspThrSer 519
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1460 TCATTCAAGAAATATGATTCAGTGTGATGCAAGCTGTGATGACAGCTCT 1509
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
520 GlyLys 521
   ::::::::::
1510 GGTAAA 1515

```

seq_name: /cgn2_6/prodata/1/lna/5D_COMB.seq:us-08-913-477-16

seq_documentation_block:

Sequence 16, Application US/08913477

Patent No. 5985285

GENERAL INFORMATION:

APPLICANT: Tibball, Richard W.

APPLICANT: Williamson, Ethel D.

APPLICANT: Leary, Sophie E.C.

APPLICANT: Bennett, Alice M.

APPLICANT: Oyston, Petra C.F.

TITLE OF INVENTION: VACCINES FOR PLAGUE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYTE P. C.

Wed Aug 23 11:46:33 2000

us-08-6999-716a-2.rn

Page 3

STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1462 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 8..1447
OS-08-913-477-16

```

alignment_scores:
  Quality: 2380.50      Length: 482
  Ratio: 5.001          Gaps: 1
Percent Similarity: 98.755      Percent Identity: 98.133

alignment_block:
US-08-659-716A-2 x US-08-913-477-16 ..

Align seg 1/1 to: US-08-913-477-16 from: 1 to: 1462

43 ASNALAALASPLEUTHRALASERHRRHRLAATHRALATHREULVALGI 59
   :::::|||||
2 AGCTGGCAGATTTCCTTACTGCAAGCACCCATGCAAGCACTTGTGTA 51
   |||||
59 UPFOALAAAGLLEHTRLEUTHRTYRLYSGLUJYALAPROILERHLEW 76
   |||||
52 ACAGAGCCGGCATCTCATCATATAAAGAAAGCGCTCCATAATACATTA 101
   |||||
76 ELASPAANGLYASNALLESPHRIGULEULEVALGLYHTRLEUTHRLEU 92
   |||||
102 TGGCAATGGAACAATCATACAGATTACTGTGTGGACGCTTACTCTT 151
   |||||
93 GLYGLYTYRILYSRNGLYTHTRSERHRSERVALASPNHETHRASPA 109

```

111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929

```

1052 GATAAAATTTATATAGTTATACAGATGAGAGATTTTAAAGCAGGC 1101
406 AGUTYITLysIleLeuGIuLysMetProGIInThrIleGlnValAspG 423
1102 AGAGTACAAAATTCGAGAAAATCCCTCAAAACCAATTCAGTGGATG 1151
423 LYSerGIuLysIleValSerIleLysAspPheLeuGIuSerLysn 439
1152 GGAGCGAGAAAATTAAGTCTCGATTAAGAGACTTCTGGAGGTGAGAT 1201
440 LysArgThrGIuLysIleLeuGIuLysnLysAsnSerIYSerTYAsnLy 456
1202 AAAAGAACCGGGCGCTGGTGAATCTGAAAAACATACACTGTATAATA 1251
456 SASpAsnAsnGIuLeuSerHisPheAlaThrThySerSerSplySera 473
1252 AGATTAATATATATATCTCACTTGGCCACCACTGCTGGTAAGTCA 1301
473 TGPLeuAsnAspLeuValSerGIuLysThrThrGlnLeuSerSply 489
1302 GGCCGCTCAAGACTGTGTTAGCCAAAACAACTCACTGCTGTATATT 1351
490 ThSerArgPheAsnSerAlaIleGlnLysAlaLeuAsnArgPheIleGln 506
1352 ACATCAGCTTTAATTCAGTATTAAGACACTGAACCGTTTCATTACAGA 1401
506 STYAspSerValMetGlnArgLeuLeuAspAspThrSerGIuLys 521
1402 ATATGATTCACTGATGCAACGCTGCTAGATGACACGCTGTGTATAA 1447
seq_name: /cgn2_6/prodata/1/lna/5D_COMB.seq:us-08-913-477-1

```

```

seq_documentation_block:
; Sequence 1, Application US/08913477
; Patent No. 5985285
; GENERAL INFORMATION:
; APPLICANT: Tibball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Leary, Sophie E.C.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Bennett, Alice M.
; TITLE OF INVENTION: VACCINES FOR PLAGUE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHIVE P.C.
; STREET: 1100 No. 5985285th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,477
; FILING DATE: 13-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00571
; FILING DATE: 13-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9505059.7
; FILING DATE: 13-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9518946.0
; FILING DATE: 15-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9524825.8
; FILING DATE: 05-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.

```

```

; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 124-599
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Yersinia pestis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..987
; US-08-913-477-1

alignment_scores:
Quality: 1641.50 Length: 328
Ratio: 5.020 Gaps: 1
Percent Similarity: 99.695 Percent Identity: 99.390

alignment_block:
US-08-699-716a-2 x US-08-913-477-1 ...
Align seg 1/1 to: US-08-913-477-1 from: 1 to: 1014

194 GlnPheMetIleArgAlaTyrGlnGlnAsnProGlnHisPheIleGlnAs 210
7 GAATTCT...ATTAGCGCTACGAGAACAAACCCACACATTTATTATGAGA 53
210 PLeuGlnLysValArgValGlnGlnLeuThrGlyHisGlySerSerVal 227
54 TCTGAAAAAGTTAGGTGGAGAACACTTACTGTCATGCTTCTTCACTTT 103
227 euGlnGlnLeuValGlnLeuValLysAspLysAsnIleAspIleSerIle 243
104 TAGAAGATTGCTTCACTTACTTCAAGATAAATAATATAGATATTTCCATT 153
244 LysTYAspProArgLysAspSerGIuValPheAlaAsnArgValIleTh 260
154 AAATATGATCCAGAAAAAGATTCGAGGTTTTTCCATATGAGTAATTAC 203
260 rAspAspIleGlnLeuLeuLysLysIleLeuAlaTyrPheLeuProGln 277
204 TGATGATATCGAATGCTCAAGAAAAATCCAGCTTATTCTTACCCGAGG 253
277 sPThrIleLeuLysGlyGlyHisTYAspAsnGlnLeuGlnAsnGlyIle 293
254 ATGCCATTCTTAAAGCGGCGCATATATACAAACCAACTGCCAAATGGCATC 303
294 LysArgValLysGlnPheLeuGlnSerSerProAsnThrGlnTrpGln 310
304 AAGCGAGTAAAGAGTTCCTGATTCATCGCCGCAATACACAATGGGAATT 353
310 uATGAlaPheMetAlaValMetHisPheSerLeuThrAlaAspArgIle 327
354 GCGGCGCTTCATGCGAGTAATGCAATTCCTTTAAACCGCGAGTCGTATCG 403
327 sPAspAspIleLeuLysValIleValAspSerMetAsnHisGlyAsp 343
404 ATGATGATATTTTAAAGTATTTCTTATTCATGATATCAATCAGGTGAT 453
344 AlaArgSerLysLeuArgGlnGlnLeuAlaGlnLeuThrAlaGlnLeu 360
454 GCGCGTAGCAAGTGGGTGAAGAAATATGCTGAGGTTTACCCCGAATTTAA 503
360 sIleTYSerValIleGlnAlaGlnIleAsnLysHisLeuSerSerSerg 377

```

```

|||||
504 GATTATTCAGTATTCAGCCGAAATTAATACCATCTGCTAGTAGTG 553
377 lYthrlleasniIleHsAsplySerlleasnlleuMetAsplyAspleu 393
|||||
554 GCACCAATAATATCCATGATTAATCCATTATCTCATGATTAATAATTTA 603
394 TyrglyYrThraspgluIuilepHelyAslaserlaIuIyrylsII 410
604 TATGGTATACAGATGAAGATTTTAAAGCCAGCCAGAGTCAAAAT 653
410 eLeugluYsMetProglInThrlIegInValAspGlySerIuYsL 427
654 TCCTCGAATAATGCTCAACACCATTCAGTGGAGGGGCGCAGAAAA 703
427 YsIleValSerlleAspPheleuglySerIuAsnlySAspArgThgIy 443
704 AATATGCTCTGATTAAGACCTTCTTCTGAAAGTGAGATTAAGAACGGG 753
444 AlaleuglyAsnleuYsAsnSerTySerTyAsnlySAspAsnAsnG 460
754 GCGTGGGTATCTGAATAAOTACTACTATATATAAGATATAATGA 803
460 uLeuSerHisPheAlaThrThrcysSerAspIysSerArgProleuAsnA 477
804 ATATCTCATCTTGCACACCTGCTCGATTAAGTCCAGCGCGCTCAACG 853
477 sPleuValSerGlnlySthrThrgInleuSerAspIleThrSerArgPhe 493
854 ACTGGTACCAAAACAAACACTGCTGCTGATTAATCACTACGCTTT 903
494 AsnSerAlaIleGluAlaleuAsnArgPheIleGlnlySAspSerVa 510
904 AATTACACTATTGAGACACTGAACCGTTTCATTAGAAATATGATTAGT 953
510 lMetGlnArgLeuLeuAspAspThrSerGlyYs 521
954 GATGCAACGCTCTGCTAGATGACACGCTGTGTA 987

seq_name: /cgn2.6/ptodata/1/lna/5D_COMB.seq:us-08-913-477-3
seq_documentation_block:
; Sequence 3, Application US/08913477
; Patent No. 5985285
; GENERAL INFORMATION:
; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Leary, Sophie E.C.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Bennett, Alice M.
; TITLE OF INVENTION: VACCINES FOR PLAQUE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P. C.
; STREET: 1100 NO. 5985285th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,477
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00571
; FILING DATE: 13-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9505059.7

```

```

; FILING DATE: 13-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9518946.0
; FILING DATE: 15-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9524825.8
; FILING DATE: 05-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 124-599
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Yersinia pestis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..987
; US-08-913-477-3

```

```

alignment_scores:
Quality: 1641.00 Length: 325
Ratio: 5.049 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.692

```

```

alignment_block:
US-08-699-716a-2 x US-08-913-477-3

```

```

Align seg 1/1 to: US-08-913-477-3 from: 1 to: 1014

```

```

197 lIeArgAlaTyrgluGlnAsnProGlnHsPheIleGluAspLeuGluY 213
|||||
13 ATTGAGGCTTACGACAAACCAACATTTATTGAGCATCTAGAAAA 62
213 sValArgValGluGlnleuThrglyHsGlySerSerValleuGluGlu 230
|||||
63 AGTTAGGCTGGAACAACCTTACTGCTGCTGCTGCTGCTGCTGCTGCTG 112
230 euValGlnleuValYsAspIysAsnIleAspIleSerIleYsTyAsp 246
|||||
113 TGGTTCAGTATGTCAAAGATTAATAATATTTCCATTAATATGAT 162
247 ProArgYsAspSerGluValPheAlaAsnArgValIleThraspAsp 263
163 CCCAGAAAATTCGAGGCTTTTCCCATAGCAATTTACTGATGATAT 212
263 eGluLeuLeuYsIleleuAlaTyrgPheleuProGluAspThrIle 280
|||||
213 CGAATGCTCAAGAAAATCTACTATTATTACCGAGAGATCCATTC 262
280 euYsGlyGlyHsTyAspAsnGlnleuGlnAsnGlyIleYsArgVal 296
|||||
263 TTAAAGCGGCTCATTTAGAACCAACTGCMAAATGGCATCAAGCAGTA 312
297 lYsGluPheleuglyuSerSerProAsnThrgInleuGluuArgAlaPh 313
313 AAAGAGTCTTGAATCATCGCCGAATPACCAATGGGAATGGGGCGT 362
313 eMetAlaValMetHisPheSerleuThralaAspArgIleAspAspI 330
|||||
363 CATGGCAGTAATGATTTCTCTTAACCGCGCATCGATCGATGANGATTA 412

```



```

123 GYASnAsnHISGlnPheThrThrLysValIleGlyLysAspSerArgAs 139
139 PHeAspIleSerProLysValAsnGlyLysAsnLeuValGlyAspAsp 136
376 TTTTGATATCTCTCCCTAAGTAAACGTCGAGAACCTGTGGGGATCCAG 425
156 aValLeuAlaThrGlySerGlnAspPhePheValArgSerIleGlySer 172
426 TCCTCTGGCTACGGCAGCCAGCATTTCTTGTCTGCTCAATTGGTCC 475
173 LysGlyGlyLysLeuAlaIleGlyLysThrAspAlaValThrValTh 189
476 AAGCGCGGTAACTGCAGCAGGTAAATACACTGATGCTTAACCGTAAC 525
189 rValSerAsnGln 193
526 CGTATCTAACCA 538

```

seq_name: /cgn2_6/ptodata/1/ina/5D_COMB.seq:US-08-913-477-10

seq_documentation_block:

```

; Sequence 10, Application US/08913477
; Patent No. 5985285
; GENERAL INFORMATION:
; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Leary, Sophie E.C.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Bennett, Alice M.
; TITLE OF INVENTION: VACCINES FOR PLAGUE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYTE P.C.
; STREET: 1100 No. 5985285th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913.477
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00571
; FILING DATE: 13-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9505059.7
; FILING DATE: 13-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9518946.0
; FILING DATE: 15-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9524825.8
; FILING DATE: 05-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 124-599
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Yersinia pestis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..454
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: 1..6
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: 536..541
; US-08-913-477-10

```

```

alignment_scores:
Quality: 755.00 Length: 151
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.013

```

alignment_block:

US-08-699-716a-2 x US-08-913-477-10 ..

Align seg 1/1 to: US-08-913-477-10 from: 1 to: 541

```

43 AsnAlaAlaAspLeuThrAlaSerThrThrAlaThrAlaThrLeuValG1 59
2 AGCTCGCAGATTACTGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 51
59 uproAlaArgIleThrLeuThrThrLysGlyLysAlaProIleThrIleM 76
52 ACCAGCCCGCATCCTATATACATATAGAGAGCGCTCCAAATTACATTA 101
76 eAspAsnGlyAsnIleAspThrGluLeuLeuValGlyThrLeuThrLeu 92
102 TGGACATATGAAACATCGATACGAATTTCTTGTGTACGCTTACTCTT 151
93 GlyGlyThrLysThrGlyThrThrSerThrSerValAsnPheThrAspAl 109
152 GGGGCTATAAAAACAGAACACTACACACTGCTTAATTACAGATGC 201
109 AlaGlyAspProMetThrLeuThrPheThrSerGlnAspGlyAsnAsnH 126
202 CGCGGTGATCCCATGTACTTAACATTACTTCTCGAGATGAAATTAACC 251
126 IsGlnPheThrThrLysValIleGlyLysAspSerArgAspPheAspIle 142
252 ACCAATTCACATAAAGATGATGGCAAGATTCTAGAAATTTGATATC 301
143 SerProLysValAsnGlyLysAsnLeuValGlyAspAspValAlaLeuAl 159
302 TCCTCCTAAGTAAACGTCAGAACCTTGTGGGAGATGAGCTGCTTGGC 351
159 aThrGlySerGlnAspPhePheValArgSerIleGlySerLysGlyL 176
352 TACGGGCACCCAGGATTTCTTGTGCTCAATTTGTTCCAAAGGGGTA 401
176 yLeuValAlaIleGlyLysThrThrAspAlaValThrValThrValSer 192
402 AACTTCAGCAGGTAAATACACTGATGTCGTAAACCGTAACCGTATCTAAC 451
193 Gln 193
452 CAA 454

```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-07-792-865D-1

seq_documentation_block:

; Sequence 1, Application US/07792865D

[illegible]

939 CATCTACCTGACTCTTCCTGTCACCT..GTTACCAATTCATCATCA 985

449 ysasnerTyrSerTyrAsnLysAspAsn 458

986 AGAATACGCGCTGTGTACAGACAGACAT 1014

seq_name: /cgn2.6/ptodata/1/lna/5D_COMB.seq:us-08-405-496A-25

seq_documentation_block:

Sequence 25, Application US/08405496A

Patent No. 5919665

GENERAL INFORMATION:

APPLICANT: WILLIAMS, JAMES A.

TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM

TITLE OF INVENTION: NEUROTOXIN

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL, LLP

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/405,496A

FILING DATE: 16-MAR-1995

CLASSIFICATION: 424

Prior APPLICATION DATA:

APPLICATION NUMBER: US 08/329,154

FILING DATE: 25-OCT-1994

Prior APPLICATION DATA:

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

Prior APPLICATION DATA:

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992

Prior APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791

FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPD-01308

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ. ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 1402 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1386

US-08-405-496A-25

alignment_scores:

Quality: 154.00

Ratio: 0.691

Percent Similarity: 46.751

Percent Identity: 19.078

alignment_block:

US-08-699-716A-2 x US-08-405-496A-25

Align seg 1/1 to: US-08-405-496A-25 from: 1 to: 1402

Length: 477

Gaps: 23

Percent Identity: 19.078

1 MetGlyHisHisHisHisHisHisHisHisHisHisSerSerGlyHis1 17
1 ATGGCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 50
17 eAspAspAspAspLysHisMetLysLysLysSerSerValIleAlaIle 34
51 C.....GAAGTCGTCATATGCTAGCATGCTGCTGCTGCTGCTGCTG 84
34 lAluPheGlyThrIleAlaThrAlaAsnAlaAlaAspLeuThrAlaSer 50
85CTGCTACCTTC 96
51 ThrThrAlaThrAlaThrLeuValGluProAlaArgIleThrLeuThrTy 67
97 ACTGAATACATCAAGACATCATCATCAATCATCATCATCATCATCATCAT 146
67 rLysGluGlyAlaProIleThrIleMetAspAsnGlyAsnIleAspThrG 84
147 CGAA.....TCCAAATCACCTGATGACCTGT 172
84 lAluLeuValGlyThrLeuThrLeuGlyGlyThrLysThrGlyThrThr 100
173 CTCGCTACGCTTCCAAATCATCATCATCAT..... 201
101 SerThrSerValAsnPheThrAspAlaAlaGlyAspProMetThrLeuTh 117
202 ...TCTAAAGTTACTTC.....GATCCGATC..... 225
117 rPheThrSerGlnAspGlyAsnAsnHisGlnPheThrThrLysValIleG 134
226GACAAAGATCAGATCCAG.....CTGTCA 250
134 lLysAspSerAlaGspPheAspIleSerProLysValAsnGlyGluAsn 150
251 ATCTGATCTTCCAAATCATCATCATCATCATCATCATCATCATCATCAT 288
151 lAluValGlyAspAspValAlaLeuAlaThrGlySerGlnAspPhePheVa 167
289 GCTATCTATACATCATCATCATCATCATCATCATCATCATCATCATCAT 338
167 lArgSerIleGlySerLysGlyGlyLysLeuAlaAlaGlyLysThrThra 184
339 CCGT.....ATCCGAATACTCA 358
184 sPAlaValThrValThrValSerAsnGlnGluPheMetIleArgAlaTyr 200
359 AC.....TCCATCTCTGACATGAAATACACATCATCATCATCATCATG 402
201 GluGlnAsnProGlnHisPheIleGluAspLeuGluLysValArgValG1 217
403 GAATAACATTC.....TGTTGGAAGATATCTGTGAATCAGTGGA 445
217 uGlnLeu.....ThrGlyHisGlySerSerValLeuGluGluLeuValG 232
446 TCATCTGACTCTGACAGACATCATCATCATCATCATCATCATCATCAT 495
232 lAluLeuValLysAspLysAsnIleAsp..... 240
496 AAATACTCTGAGATGATCAATCATCTGACTACATCAATGCTGATCTT 545
241 lIleSerIleLysThrAspProArgLysAspSerGluValAlaPheAlaAsn 257
546 CGTACCATCAACCAATGCTGTGAATACTCCAATATCTACATCAACG 595
257 rValIleThrAspAspIleGluLeuLeuLysLysIleLeuAlaTyrPhe 273
596 GCCCTCTGATCGACAG.....AAACGATC..... 621
274 lLeuProGluAspThrIleLeuLysGlyGlyHisThrAspAsnGlnLeuG1 290
622TCCAAATCTGGG 632

```

290      nasglnyllelysarvalylsgluipheleuglserseprosnrhg 307
      |||||
653      TAAC.....ATCCAGCTTCTAATATCA 658
307      intrpgluu.....Argalaphelelavalmet 317
      :|||:
659      TGTTCMAACGTGACGGTGTGCTGACACTACCGCTACATGTGATCAA 708
      :|||:
318      HspPheSerLeuThrAlaAspArgIleAspAspArgIleLeuValII 334
      :|||:
709      TACTTCATCTGTTTCGACAAAGAACTGACCAAAAGAAATCAAAAGCCT 758
      :|||:
334      evalasperketasnhshsglyaspalargserlyseuarglug 351
      :|||:
759      GTACGACACCAAGTCCAACTTGTGT.....ATCCTGAAGAGCT 786
      :|||:
351      luLeuAlaGluLeuThrIaGluLeuLysIleTyrSerValIIleGlnAla 367
      :|||:
797      TGTGGGGTACTACTCTGACAGTACGACAAACCGTACTACATG..... 837
      :|||:
368      GluIleasnLysHsLeuSerSerSergIlyThrIleasnIleHsasp.. 383
      :|||:
838      .....CTGAATCTGTACGATCC 854
384      ...LysSerIleasnLeuMetAspLysAsnLeuTyrGlyTyrThrAspG 399
      :|||:
855      GACAAATACCTGTGACGCTCAACAAATGATGATATCCGGCGTTAC..... 897
      :|||:
399      luGluIlePheLysAlaSerAlaGluTyrLysIleLeuGluLysMetPro 415
      :|||:
897      ..... 897
416      GluThrThrIleGlnValAspGlySerGluLysIleValSerIleLy 432
      :|||:
898      .....ATGTAACCTGAAGGTCGCGGTGCTGTCTGTAATGACTACCA 938
      :|||:
432      sAspPheLeuLysSerGluAsnLysArgThyGlyAlaLeuGluLysLeuL 449
      :|||:
939      CATCTACCTGACACTTTCCTCCGTACCCG...GGTACCAAAATCATCATCA 985
      :|||:
449      ysAsnSerTyrSerTyrAsnLysAspAsn 458
      :|||:
986      AGAAATACCGCTGTGTTACCAAGACAACT 1014
seq_name: /cgn2.6/ptodata/1/lna/6_COMB.seq:US-08-870-370-7
seq_documentation block:
: Sequence 7, Application US/08870370
: Patent No. 6060242
: GENERAL INFORMATION:
: APPLICANT: Eileen Nie and Yuan Min Wu
: TITLE OF INVENTION: PNA Diagnostic Methods
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Caesar, Rivise, Bernstein, Cohen &
: ADDRESSEE: Pokotilow, Ltd.
: STREET: 12th Floor, 7 Penn Center, 1635 Market
: STREET: Street
: CITY: Philadelphia
: STATE: Pennsylvania
: COUNTRY: U.S.A.
: ZIP: 19103-2212
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Tenet, David M.
: REGISTRATION NUMBER: 37,054

```

```

REFERENCE/DOCKET NUMBER: E1047/20001
TELEPHONE: 215-567-2010
TELEFAX: 215-751-1142
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 bases
TYPE: nucleotide
STRANDEDNESS: double-stranded
TOPOLOGY: linear
US-08-870-370-7

alignment_scores:
Quality: 146.00      Length: 22
Ratio: 6.636         Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-08-699-716A-2 x US-08-870-370-7  ..

Align seg 1/1   to: US-08-870-370-7   from: 1   to: 375

1 MetGlyASHSHSHSHSHSHSHSHSHSHSHSHSHSHSerglynsil 17
|||||
89 ATGGCGCATCATCATCATCATCATCATCATCATCATCAGCAGCGGCATAT 138
|||
139 CGACGACGACGACGAAG 154

seq_name: /cgn2_6/pctodata/1/lna/6_COMB.seq:US-08-870-370-8

seg_documentation_block:
Sequence 8, Application US/08870370
Patent No. 6080242
GENERAL INFORMATION:
APPLICANT: Eileen Nie and Yuan Min Wu
TITLE OF INVENTION: PNA Diagnostic Methods
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Caesar, Rivise, Bernstein, Cohen &
ADDRESSER: Pokocillow, Ltd.
STREET: 12th Floor, 7 Penn Center, 1635 Market
Street, Street
City: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103-2212

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/870,370
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Tener, David M.
REGISTRATION NUMBER: 37,054
REFERENCE/DOCKET NUMBER: E1047/20001
TELEPHONE: 215-567-2010
TELEFAX: 215-751-1142
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 bases
TYPE: nucleotide
STRANDEDNESS: double-stranded
TOPOLOGY: linear
US-08-870-370-8

```



```

1368 TGCAGGCTGCTTAAAGAAAGTGCCTACTAATGCAAC..... 1411
108 AspaLaalaglyasprometylLeuThrPheThrSerGlnaspGlyAs 124
1412 .....GATCTTCATCAAGCTCTCACAC 1434
124 nasnHisGlnPheThrThrLysValIleGlyLysaspSerArgasp... 139
1435 CGTCTTACATCAACAAACATAGTTGAGCAAGATGCAAGACCATG 1484
140 .....PheaspIleSerProLysValasnGlyLysasnLeu 151
1485 GACAAACAAATAATTTGTTAAATGGCAATGGAATTTGCTGATTTG 1534
152 ValGlyaspAspValValLeuAlaThrGlySerGlnaspPheValar 168
1535 GCAGCGAGTGAACATCAACAGATCGGCTCGGAGATTAAGGCGCA 1584
168 gserIleGly...SerLysGlyLysLeuAlaAlaGlyLysThr 184
1585 AGAAGCTGGCTAATAAACAATGCTGTACACTAGCGCTTTATCA 1634
184 spAlaValThrValThrValSerasnGlnIuphMetIleArgAlaTyr 200
1635 ACGCAGTC...GTGATCATCTTCAACAT.....ATACCTTACAGA 1672
201 GlnGlnaspProGlnHisPheIleGlnasp.....LeuGlyLys 213
1673 GAATCTAAGCTAACACAGATGCTACAGACTTTAGTGGTATGACGAA 1722
213 svaIArgValGlnLeuThrGlnHisGlySerSerValLeuGlnGlu 230
1723 AACATGCAATTTATCGCAACTATATGACCTGCAAAATATTCATGAA 1772
230 euValGlnLeuValLysaspLysasnIleaspIleSerIleLysTyr 246
1773 CTGCAACTAGCTGATGATGATGCAAGAGAGCAATCATTAAGAAACT 1822
247 ProArgLysaspSerGluValPheAlaAsnArgValIleThrAsp 263
1823 CCACAGATTAATCACTCT..... 1840
263 eglLeuLeuLysLysIleLeuAlaTyrPheLeuProGlnaspThrIle 280
1841 .....TTATCGAAGATATCACTGTC 1859
280 euLysGlyGlnHisTyrAspAsnGlnLeuGlnasnGlyLysArgVal 296
1860 TCAAA.....GACTACATTCAGAG..ATTGAAATA 1891
297 LysGlnPheLeuGlnSerSerProAsnThrGlnTyrPheLeuArgAla 313
1892 AGAATGATTTGAAAAATTCAGAAACAAACNA..... 1924
313 emeAlaValMetHisPheSerLeuThrAlaAspArgIleasp...Asp 329
1925 .....GCTATATTATACCAAGATGAGTGGACCTTACG 1961
329 spIleLeuLysValIleValaspSerMetasn.....HisIleGly 342
1962 AGACCAATTCATCTGATTCATGAGCAAAATCTAAATAATCATATAC... 2008
343 AspaLaalagSerLysLeuArgGlnLeuAlaGlnLeuThrAla.... 357
2009 .....CTGCGAGAACAAATTAATAATTCAGAAACAACTA 2043
358 .....GlnLeuLysIleTyrSerValIleGlnAlaGlnLeuLys 372
2044 CCGAACCAATTCATGATATCAATAATCTTTTACAGTCAAGAAAGAAAC 2093
372 LysLeuSerSerSerGlyThrIleasn..... 380
2094 TAATTCGCAATATACAGAAATTTAATGTCATTTTCTACTTTACTCG 2143

```

```

381 .....IleHisaspLysSerIleasnLeuMetaspLysasnLe 393
2144 GAATCCAAAAAATTCACACTACTAATCTGAACTAATGAA...GT 2190
393 uTyrGlyTyrThrAspGlnGluIlePheLysAlaSerAlaGlnTyrLys 410
2191 CATACAAACAGAGATTTTTCACACTAGAAATTCACAAAACAGTATATA 2240
410 LeuGlnLysMetProGlnThrThrIleGlnValaspGlySerGlyLys 426
2241 CGAACCGAAGCATG.....CAATTAAATCTCTCAACAA 2275
427 LysIleValSerIleLysaspPheLeuGlySerGlnasn..... 439
2276 GTTTACAGACTTTGAAACACTTACAGGCTCTTAAATATATTAATC 2325
440 LysArgThrGlyAlaLeuGlnLysasnLeuLysasnSerTyrSerTyr 456
2326 TAATGTTCGAAAGTTATCAAGGCGTACCGAAGACTAACCAGGAACG 2375
456 LysAspAsnAsnGlnLeuSerHisPheAlaThr..... 466
2376 TAATACCAATTAAGCGCAACACGATTCCTCTCAATGCTATTAAC 2425
467 ...ThrcysSeraspLysSerArgProLeuasnaspLeuValSerGln 482
2426 ATTTACTACTACTATGATGATGATGATGATGATGATGATGATGAT 2475
482 sThrThrGlnLeu.....SeraspIleThrSerArgPheasn 495
2476 TTCACCTCTCAATGAATATTTCAAGAGTGTACTCTCTCACTATGCT 2524
495 eraIaIleGlnAlaLeuAsnArgPheIleGlnLysTyrAspSerVal 511
2525 .....AAGATTTCATGAATCTACCAATCACAATCACAATTTCTA 2566
512 GlnArgLeuLeuaspasp 517
2567 AAAAATTTACAAAACGAT 2584

```

seq_name: /cgn2.6/plodata/1/ina/5C_COMB.seq:US-08-376-843-23

```

seq_documentation_block:
; Sequence 23, Application US/08376843
; Patent No. 5846764
; GENERAL INFORMATION:
; APPLICANT: Demaggio, Anthony J.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/376,843
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5846764and, Greta E.

```


1773 CTGCAAGTACGCTAGAAATATGCAACGAGAGCCAAATCAATTAGAATACT 1822

2525AAAGTTTGATGAAATCTACCAATCACATCAACAAATTTCCTA 2566

2525AAAGATTGATGAATCTACCAATCACATCAACAATTCTA 2566


```
906 AAGAGAACTCAGATGTGATAGCTCTATGGAAGTCCAAACTGGAGACTGCC 955
501 AsnArgPheIleGlnLysTyrAspSerValMetGlnArgLeu 514
    ||| .....|||.....|||
956 .....ATCGCATCCCAACACACAGCGCATGGAAGACTG 988
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

[illegible]

```

81   lileapThnglu.....leulevalGlyThrleuthenrlenglygtyt 95
      |||||
583 ACCTTACCGCAATCGCAGTCGACGCCAGCCGCTGTGAAGTGACAGGGCTTC 632
      |||||
95 rlystrhglythrThrSerTherSeValaAanPhetrhspALaAla1ya 112
      |||||
633 TGGTACTGGAAACCATGACGAAGCGATTGTGAAAATAATGACGCCACTGCAC 682
      |||||
112 sspmroetryrLeutrhpe 118
      ::|||
683 ATCGCTCGCATTTGTTCTTC 702

seq_name: gb_estl4:AA941922

seq_documentation_block:
LOCUS       AA941922                713 bp    mRNA           EST          25-NOV-1998
DEFINITION  LD37265.5prline LD Drosophila melanogaster embryo POT2 Drosophila
            melanogaster cDNA clone LD37265 5prime, mRNA sequence.
ACCESSION   AA941922
VERSION     AA941922.1  GI:3102473
KEYWORDS    EST.
SOURCE      fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 713)
AUTHORS    Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
            Brokstein,P., Lewis,S. and Rubin,G.M.
TITLE      BDGP/NHMT Drosophila EST Project
JOURNAL    Unpublished (1997)
COMMENT    On Jan 19, 1998 this sequence version replaced gi:12284650.
            Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 LSA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
FEATURES             Location/Qualifiers
     source          1..713
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_id="LD37265"
                     /cdate="19980119"
                     /library="BDGP"
                     /method="cDNA"
                     /notes="Embryo; Vector: pOT2; Site: 1; EcoRI; Site 2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2."
BASE COUNT        167 a         215 c         194 g         137 t
ORIGIN
alignment_scores:
Quality: 115.00 Length: 140
Ratio: 1.691 Gaps: 6
Percent Similarity: 48.571 Percent Identity: 29.286

alignment_block:
US-08-699-716A-2 x AA941922 ..

Align seg 1/1 to: AA941922 from: 1 to: 713

2 GYHNISHISHISHISHISHISHISHISHISHISerSergLYNStleas 18
|||||
303 GGCGCATCATCGGCCACCACCATCACACCCACAAGAAGTGCACAGCGGA 352
|||||
18 pasparaparplynshmet.....lysl 26
|||||

```

353 CGAGACGAGACACACAACCTTAAGTTGCCAGCAGCAACTCTGACGAAM 402

26 ystleserSerrValileAlaleupheUlyhrillealatrha 42
|||||
:::
403 AGCGGGGATGCCTGCTGTGCTCCACTGACCTTACCCTATAACGC 452
453 ATGGGTGCC.....CTGAGTCGCACCTGCTTAACGAGACCACTTCAG 486
59 UProlAargilEthrLeuthrYrrysgluglYala.....P 72
::|||
497 TCCAGCCAGATTTC.....TACAAGCCCACTGCCGAAAGAACGCA 537
72 rollethrllwEt.....Aspasnglyasn 80
|||
538 AGATCAAACGATGTCATCATGAGTTCAGATTCCAGAAAGTATCATCAC 587
81 lileaspthglu.....leulevaliglythLeuthrleuclYglyTY 95
|||||
588 ACCTTGACCGCAATCGCACCTGCAGCCAGCCTCTAAGTAGACGCGTTC 637
95 rlystrnglyThrsErThrserValanPhetrnAspalAlagLYA 112
|||||
638 TGCTACTCGAGACCATACGACGAGGATTGTGAANAATTGACGCCACTGAC 687
112 spProWetyrLeuthrPhe 118
::||| |||
688 ATGCTGCTCATTTGTCTCTC 707

seq_name: gp_est14:AA942159

seq_documentation_block:

LOCUS AA942159 727 bp mRNA EST 25-NOV-1998

DEFINITION LD6225.5prime LD Drosophila melanogaster embryo POT2 Drosophila melanogaster cDNA clone LD6225 5prime, mRNA sequence.

ACCESSION AA942159

VERSION AA942159.1 GI:3101782

KEYWORDS EST.

SOURCE fruit fly.
Drosophila melanogaster

ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 727)
Harvey,D., Hong,L., Evans-HoM,M., Pendleton,J., Su,C., Brockstein,P., Lewis,S. and Rubin,G.M.
BDGP/HMT Drosophila EST Project
Unpublished (1997)

TITLE JOURNAL COMMENT

On Jan 19, 1998 this sequence version replaced gi:2284870.
Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSJ, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 262 row: C column: 1
High quality sequence stop: 698.

FEATURES

location/Qualifiers

1..727

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="LD6225"

/clone_lib="LD Drosophila melanogaster embryo POT2"

/sex="male and female"

/dev_stage="0 to 24 hours mixed stage embryonic"

/lab_host="XLI Blue"

/note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2."

BASE COUNT 173 a 224 c 193 g 137 t

ORIGIN


```

21 AsplyshtisMethylseriservValIleAlaIleAlaLeuPheG1 37
231 .....GCCACGCTATACACACACACACACAC 259
37 YThrIleAlaThrAlaAsnAlaAlaAspLeuThrIleAspThrAlaT 54
260 CACCATACCC.....ACGCCATACACACCGGTA 288
54 hralaThrLeuValGluProAlaArgIleThrLeuThrTyrLysGluGly 70
289 CGGGACCGGTATTCACACCGCCACCGCTACCGTCACT..... 326
71 AlaprolleThrIleMetaspasnGlysnIleAspThrGluLeuVal 87
326 ..... 326
87 lGlyThrLeuThrLeuGlyGlyTyrLysThrGlyThrThrSerThrServ 104
327 ....ACGACACACATA.....TACACCGCCACACGATGTATCTCT 363
104 aAlaAspThrAspAlaAlaGlyAspPrometTyr...LeuThrPheThr 119
364 CCTCTCTCCGACGCGCGCGCGCCGACCGCGGACTGAGTTCATT 413
120 SerGlnAsp..... 122
414 GACAGTATATTAACGCGCGCTCACTGTCTCATCTCCGTCGATCG 463
123 .....GlyAsnAsnHisGlnPheThrThrLysV 132
464 CATCTTCATGACTGCTCCGAGCGAGATCAATCAAAACCCCTG 513
132 aIleGlyLysAspSerArgAspPheAspIleSerProLysValasnGly 148
514 TA.....TCTGCCAGTCTCACGATCGTT..... 536
149 GluAsnLeuValIleGlyAspAspVal...ValLeuAlaThrGlySerGlnAs 164
537 GATATCTCTCTTCACGATATCGCTACTGAGAGAAGGCTCTCTGA 586
164 pPhePheValArgSerIleGlySerLysGlyGlyLysLeuAlaGlyL 181
587 GGGCATGCTCTCACT.....GGCCCAACG 612
181 ySTyrThrAspAlaValThrVal 188
613 AGCCCAACCGCAACGCGTA 635
seq_name: gb_est28:AU003803

```

```

seq_documentation_block:
LOCUS AU003803 753 bp mRNA EST 19-JAN-1999
DEFINITION AU003803 Bombyx mori p50(Dalzo) Bombyx mori cDNA clone ws00597,
ACCESSION AU003803
VERSION AU003803.1 GI:4161174
KEYWORDS EST.
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Bombycoidea; Bombycidae; Bombyx.
1 (bases 1 to 753)
AUTHORS Mita,K., Moriyama,M., Shinada,T., Okano,K. and Maeda,S.
TITLE Establishment of cDNA database of Bombyx mori
JOURNAL Unpublished (1999)
COMMENT On May 11, 1999 this sequence version replaced gi:477651.
Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmita@nirs.go.jp
PROJECT - 'CREST project by JST'.

```

```

FEATURES
    source
        1..753
        /organism="Bombyx mori"
        /strain="p50(Dalzo)"
        /db_xref="taxon:7091"
        /clone="ws00597"
        /clone_1fb="Bombyx mori p50(Dalzo)"
BASE COUNT      243 a      177 c      220 g      113 t
ORIGIN
alignment_scores:
    Quality: 108.50      Length: 311
    Ratio: 0.719      Gaps: 11
    Percent Similarity: 48.553      Percent Identity: 18.650
alignment_block:
US-08-699-716A-2 x AU003803
Align seg 1/1 to: AU003803 from: 1 to: 753
215 ArgValGluGlnLeuThrGlyHisGlySerSerValLeuGluGluLeuVal 231
|||.....|||
6 CGAGCCGATCAAGCTTTAAAGAGAGCCAGAGCTGTATGACGAACTTT 55
231 lGlnLeuValLysAspLysAsnIleAspIle.....SerIleLysTyrA 246
|||.....|||
56 AGTCTGTACGCGGAGGTCAACACACGCTGCTGCCGACATCAAGCTCA 105
246 sProAlaGlyAspSerGluValAlaPheAlaAsnArgValIleThrAspAsp 262
|||.....|||
106 ACAAGCTTCACGAAGACTCAATGAGATGATGAAAGATGATGAGAAA 155
263 lIleGluLeuLysLysIleLeuAlaTyrPheLeuProGluAspThrI 279
|||.....|||
156 TCGGCAGAGCTGACCGCTTGCATGCTGAG.....AACGAGACACCT 199
279 eLeuLysGlyGlyHisTyrAspAsnGlnLeuGlnAsnGlyLleLysArgV 296
|||.....|||
200 T.....CACAAATTAAAGAGGCGCATAGGCGCG 228
296 aLysGluPheLeuGluSerSerProAsnThrGlnTProGluLeuArgAla 312
|||.....|||
229 GTCGGAGCTTGTCTGGAACAGGTACAGACGAGGAGAACTCAAGCAT 278
313 PheMetAlaValMetHisPheSerLeuThrAlaAspArgIleAspAsp 329
|||.....|||
279 CTCTCGCTAGCTG.....GATGA 298
329 pIleLeuLysValIleValAspSerMetAsnHisGlyAspAlaArgS 346
|||.....|||
299 GCTG.....CAGCGCAAGCTTAAA 318
346 eLysLysLeuArgGluGluLeuAlaGluLeuThrAlaGluLeuLysIleTyr 362
|||.....|||
319 ACGATGTGGAAGCTCAGAAAGCCAGTTGACAGCGCAATGAAATTATAC 368
363 serValIleGlnAlaGluIleAsnLysHisLeuSerSerGlyThrI 379
|||.....|||
369 AAGACTTTAAA...GACTTCAGATGATCAAGTACCCAGTGC..... 407
379 eAsnIleHisAspLysSerIleAsnLeuMetAspLysAsnLeuTyrGlyT 396
|||.....|||
408 .....CACAGCAAGCAGATC..... 422
396 YThrAspGluGluIlePheLysAlaSerAlaGluTyrLysIleLeuGlu 412
|||.....|||
422 ..... 422
413 LysMetProGlnThrThrIleGlnValAspGlySerGluLysLysIleVal 429
|||.....|||
423 .....CAAGCAGATGAGAGTCCGCGCTCCAAAGAGAGTAC 463

```



```

REFERENCE
AUTHORS
1 (bases 1 to 453)
Quatran,R., Bashlades,S., Cove,D., Cumling,A., Knight,C.,
Clifton,S., Marra,M., Hillier,L., Pepe,D., Matlin,J., Wylie,T.,
Underwood,K., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Steptoe,M., Gibbons,M., Harvey,N., Ritter,E.,
Jackson,Y., McCann,R., Waterston,R. and Wilson,R.
TITLE
Leeds/Mash v Moss EST Project
JOURNAL
Unpublished (1999)
COMMENT
On Nov 2, 1998 this sequence version replaced g1:3828423.
Contact: Ralph Quatran
Leeds/Mash v Moss EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Libraries were constructed by Dr. Stavros Bashlades as part of the
Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and
Washington Univ. in St. Louis (USA) DNA sequencing by: Washington
University Genome Sequencing Center For information on obtaining a
clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
Seq primer: -40RP from Gldco
High quality sequence stop: 420.

FEATURES
SOURCE
1. 453
/location/Qualifiers
/organism="Physcomitrella patens"
/db_xref="taxon:3218"
/clone_pep_source_id="PU111110"
/clone_id="Moss EST library PPU"
/tissue_type="protonemata: 7 day old tissue
ammonium-grown"
/lab_host="DH10B"
/note="Vector: bluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Construction of the cDNA library was carried out
using Stratagene 'UnizAP - cDNA synthesis kit'. cDNA
was constructed using an oligo dt primer/linker that
contains a XhoI site within it. Following ds cDNA
synthesis, EcoRI adapters were ligated to the blunt ends
and sample was digested with XhoI. The result is cDNA
with an EcoRI sticky end on one side and a XhoI sticky
end on the other. This cDNA was ligated directionally in
UnizAP arms. The vector is designed containing the
bluescript sequence as well as lambda DNA and cDNA is
cloned within this bluescript sequence. The vector was
then packaged using Gold gigapackaging extracts. Library
was grown in XL1blue MRF' cells and amplified. The library
was excised by mass excision using Stratagene 'Mass
excision kit' that uses exsist as a helper phage that
releases the bluescript sequence and circularizes it as
single stranded plasmids that are then packaged (by helper
phage) and secreted out of the host cell as phagendms.
SOLR cells were transformed with phagendms and the library
was plated out on LB-amp plates to select for
transformants. Approximately 1,000,000 colonies were grown
and recovered. The double stranded plasmid library was
recovered by using Qulagen Midi prep kit. 2 micro grams of
each library were used to transform DH10B cells by
electroporation."

BASE COUNT
123 a 165 c 98 g 67 t
ORIGIN
alignment_scores:
Quality: 105.50 Length: 135
Ratio: 1.575 Gaps: 10
Percent Similarity: 49.630 Percent Identity: 32.593

alignment_block:
US-08-699-716A-2 x AWS09929
Align seg 1/1 to: AWS09929 from: 1 to: 453
3 HSHSHSHSHSHSHS1S.....7

```

```

1 CACATCATCACCACCAACAGCAAGCTCCAGCGTGGCCGATGTCTACCA 50
|||||
8 ..... HSHSHSHSHSHSISerGlyHisT 17
|||||
51 TAGCAGCACCCATTGGGCACCGTCATCAGGTCAATAGTAGCCATACA 100
|||||
17 leasp...Aspasplyshsmetyls..lleserValIleAl 32
|||
101 GGAGCACTGGGGCGATGCCACTCTTAAGATGTCATGACACCATCAT 150
|||||
32 attlealeuphegylThrIleAthrAlaAsnAlaAlaspleuthra 49
::: ||||| |||||
151 ACCCGGGAACATGCACACCATR....GTCATCTCAGCAGACATGGCCA 194
|||||
49 lserThr.....AlathrAlatrHleuValGlu 59
|||||
195 GAAGCACCCTCTCGTCTCCGAGCGGTACACAGCAGCATGTC... 241
|||||
60 proAlatrgylThrLeuThrTyrlYrselulYalatrIethr...Il 75
::: |||||
242 ACCACAGGATCACCAGGTC.....ATCACCATGT 273
|||||
75 emetaspansnglyAsnIleaspThrgluLeuLeuValglYthrleuthrL 92
|||||
274 CATGCACAC.....ATCACGCGCTCGCATGGCTCCACGACACCA 314
|||||
92 euglglyTyrlYs.....ThrglyThrSerThr 102
::: |||||
315 CTCACGCTCTCCGAGCGACATCAGCAGTCACAGGGGACACAGACACC 364
|||||
103 Ser 103
:::
365 ACA 367

seq_name: gb_est7.AA439056

seq_documentation_block:
LOCUS AA439056 499 bp mRNA EST 28-NOV-1998
DEFINITION LD13501.5sprime LD Drosophila melanogaster embryo Bluescript
Drosophila melanogaster cDNA clone LD13501 5prime, mRNA sequence.
ACCESSION AA439056
VERSION AA439056
KEYWORDS EST.
SOURCE fruit fly
ORGANISM Drosophila melanogaster
Eukaryote; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 499)
Harvey,D., Hong,L., Evans-Holim,M., Pendleton,J., Su,C.,
Broksfeld,P., Lewis,S. and Rubin,G.M.
BDGP/HMT Drosophila EST Project
Unpublished (1997)
CONTACT: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 USA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: j35 row: A column: 1
High quality sequence stop: 427.
Location/Qualifiers
1..499
/organism="Drosophila melanogaster"
/db_xref="BDGP_EST:BDcldn12732"
/db_xref="taxon:7227"
/cclone="LD13501"
/xclone_lib="LD Drosophila melanogaster embryo Bluescript"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"
/note="Organ: embryo; Vector: Bluescript SK; Site:1;"
```

```

BASE COUNT      107 a   154 c   145 g    93 t
ORIGIN
alignment_scores:
    Quality: 105.50          Length: 82
    Ratio: 2.638            Gaps: 3
Percent Similarity: 48.780     Percent Identity: 34.146

alignment_block:
US-08-699-716A-2 x AA439056 ..

Align seg 1/1 to: AA439056 from: 1 to: 499

2 GLYHISHSHSHSHSHSHSHSHSHSHSHSHSHSerglyHisIleas 18
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 GGCGACATCAGGCCACCACCATCCACCACTGACCAAGTGCACAGCTGA 349
18 pAspaAspLysHisMetLysLysIleSerSerValIleAlaIleAla 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
350 CGGACTGGACACACAACTTAAGTTCCA..... 380.
35 eupheLythrIleAlaThrAlaAsnAlaAlaSplLeuThrAlaSerThr 51
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
381 .....GCCAGCGACTCTGACGAAAAGGCG 404
52 ThrAlaThrAlaThrLeuValGIuProAlaArgIleThrLeuThrTyrl 68
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
405 GAGCGATCTGCTGTCTGCTGCCA.....CTGACCTTCCTATTACG 448
68 sgluGIyAlaProIleThrIleMetAspAnslYasnIleAsprThr 83
: ||||| :|||::: ||| |||||
449 CATGGTGCC.....CTGAGTTCGACTCGCTTAACGAGCACACC 488

seq_name: gb_est28:AJ388692

seq_documentation_block:
LOCUS AJ388692 503 bp mRNA EST 21-MAR-2000
DEFINITION Medicago truncatula R108 Medicago truncatula cDNA clone
METHOD similar to glycine-rich RNA binding protein, mRNA sequence.
ACCESSION AJ388692
VERSION AJ388692.1 GI:6603958
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales;
Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 503)
AUTHORS Gyorgyev J., Vaubert, D., Jimenez-Zurdo, J.I., Charon, C.,
Troussard, L., Kondorosi, A. and Kondorosi, E.
TITLE Analysis of Medicago truncatula nodule expressed sequence tags
JOURNAL Mol. Plant Microbe Interact. 13 (1), 62-71 (2000)
MEDLINE 20120379
COMMENT On Apr 3, 1998 this sequence version replaced gi:3018972.
Contact: Gyorgyev J
Institut des Sciences Vegetales
Centre National de la Recherche Scientifique (CNRS)
Av. de la Terrasse Bat.23, Gli-sur-Yvette, 91198, FRANCE
Additional information about these EST clones can be downloaded
from http://www.cnrs-gif.fr/isv/AK/index.html. (URL provided by the
author). Location/Qualifiers
    1..503
       /organism="Medicago truncatula".
       /strain="R108"
       /db_xref="taxon:3880"
       /clone="MCN0051"
       /clone_1tb="Medicago truncatula R108"
       /tissue_type=" symbiotic root nodule"

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 14:58:47 ; Search time 1417.04 Seconds
(without alignments)
1972.304 Million cell updates/sec

Title: US-08-699-716a-1
Perfect score: 1566
Sequence: 1 ATGGGCGATCATCATCATCA.....ATGACACGTCGTGTAATGA 1566

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 972840 segs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_cm:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pl4:*
11: gb_pl5:*
12: gb_pl6:*
13: gb_pl7:*
14: gb_pl8:*
15: gb_pl9:*
16: gb_pl10:*
17: gb_pl11:*
18: gb_pl12:*
19: gb_pl13:*
20: gb_pl14:*
21: gb_pl15:*
22: gb_pl16:*
23: gb_pl17:*
24: gb_pl18:*
25: gb_pl19:*
26: gb_pl20:*
27: gb_pl21:*
28: gb_pl22:*
29: gb_pl23:*
30: gb_pl24:*
31: gb_pl25:*
32: gb_pl26:*
33: gb_pl27:*
34: gb_pl28:*
35: gb_pl29:*
36: gb_pl30:*
37: gb_pl31:*
38: gb_pl32:*
39: gb_pl33:*
40: gb_pl34:*
41: gb_pl35:*
42: gb_pl36:*
43: gb_pl37:*

44: gb_htg7:*
45: em_htg7:*
46: em_htg8:*
47: em_htg9:*
48: em_htg10:*
49: gb_pl13:*
50: gb_pl14:*
51: gb_pl15:*
52: gb_pl16:*
53: gb_pl17:*
54: gb_pl18:*
55: gb_pl19:*
56: gb_pl20:*
57: gb_pl21:*
58: gb_pl22:*
59: gb_pl23:*
60: gb_pl24:*
61: gb_pl25:*
62: gb_pl26:*
63: gb_pl27:*
64: gb_pl28:*
65: gb_pl29:*
66: gb_pl30:*
67: gb_pl31:*
68: gb_pl32:*
69: gb_pl33:*
70: gb_pl34:*
71: gb_pl35:*
72: gb_pl36:*
73: gb_pl37:*
74: gb_pl38:*
75: gb_pl39:*
76: gb_pl40:*
77: gb_pl41:*
78: gb_pl42:*
79: gb_pl43:*
80: gb_pl44:*
81: gb_pl45:*
82: gb_pl46:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	DB ID	Description
1	1473.6	94.1	1530	5	A56814
2	1410.6	90.1	1462	5	A56808
3	979.4	62.5	2100	1	YEPICR
4	979.4	62.5	70305	2	YPCDI
5	979.4	62.5	70504	2	AF053946
6	979.4	62.5	70559	2	AF074612
7	976.8	62.4	1014	5	A46411
8	976.8	62.4	1014	5	A56793
9	974.8	62.2	1014	5	A46413
10	974.8	62.2	1014	5	A56795
11	974.6	62.2	981	2	AF167310
12	967.4	61.8	975	2	AF167309
13	957	61.1	981	1	YEPICR
14	953.8	60.9	2201	1	YEPICR
15	927.4	59.2	69673	2	AF102990
16	916.2	58.5	975	1	YEPICR
17	914.6	58.4	975	1	YEPICR
18	913	58.3	975	1	YEPICR
19	911.4	58.2	975	1	YEPICR
20	890.4	56.9	1002	1	YEPICR
21	882.4	56.3	1002	1	YEPICR
22	516.8	33.0	547	5	A56812
23	516.8	33.0	5383	1	YPCAF
24	516.8	33.0	96210	2	YPCAF

```

c 25 516.8 33.0 100994 2 AF053947 AF053947 Yersinia
26 516.8 33.0 100990 2 AF074611 AF074611 Yersinia
27 217.6 13.9 240 1 S38727 S38727 IcrGVH Oper
28 75 4.8 4168 2 AF010149 AF010149 Pseudomon
29 74 4.7 660 14 AF081364 AF081364 Synthetic
30 53.4 3.4 521 10 S75264 S75264 Wtl-Wilms'
31 53.4 3.4 1402 5 AR000030 AR000030 Sequence
32 48.2 3.1 85779 8 SCE011856 SCE011856 Saccharom
33 48.2 3.1 7218 5 I66494 I66494 Sequence 14
34 47.8 3.1 252102 60 AC005140 AC005140 Plasmodu
35 47.8 3.1 256172 41 AC005139 AC005139 Plasmodu
36 46.8 3.0 110000 31 PFMAL1P1_0 PFMAL1P1_0 Plasmodu
37 46.8 3.0 293431 31 PFMAL1P4 PFMAL1P4 Plasmodu
38 46.4 3.0 1158 33 AF044866 AF044866 Phoebs s
39 45.8 2.9 157340 43 AC007926 AC007926 Trypanoso
40 45.4 2.9 282806 60 AC006279 AC006279 Plasmodu
41 45.2 2.9 21202 19 PFSC04088 PFSC04088 Plasmodu
42 45.2 2.9 224448 31 PFMAL1P4 PFMAL1P4 Plasmodu
43 44.8 2.9 1978 33 DDIDDKA DDIDDKA D.discoidu
44 44.8 2.9 12029 34 AE001426 AE001426 Plasmodu
45 44.4 2.8 321003 31 PFMAL1P3 PFMAL1P3 Plasmodu

```

ALIGNMENTS

```

RESULT 1
LOCUS A56814 1530 bp DNA PAT 03-MAR-1998
DEFINITION Sequence 22 from Patent WO9628551.
ACCESSION A56814
VERSION A56814.1 GI:3712827
KEYWORDS
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.

```

```

REFERENCE 1 (bases 1 to 1530)
AUTHORS Tibball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,
TITLE Bennett and Alice,M.
JOURNAL VACCINES FOR PLAGUE
COMMENT Patent: WO 9628551-A 22 19-SEP-1996;
SECUR DEFENCE (GB)
Other publication ZA 9602036 960716
Other publication AU 4951196 961002.
FEATURES
location/Qualifiers
source 1..1530

```

```

CDS
/organism="Yersinia pestis"
/db_xref="taxon:632."
13..1518
/feature="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CA03423.1"
/db_xref="GI:3712828"
/transl_table=11

```

```

BASE COUNT 500 a 296 c 311 g 423 t
ORIGIN

```

```

Query Match 94.1%; Score 1473.6; DB 5; Length 1530;
Best local Similarity 99.1%; Pred. No. 0;
Matches 1495; Conservative 0; Mismatches 4; Indels 9; Gaps 1;
QY 68 ATATGAAAAAATCAGTCCGTTATCCGCAATGATTTGGAACATATGCAACTGCTA 127

```

```

|||||
Db 11 ATATGAAAAAATCAGTCCGTTATCCGCAATGATTTGGAACATATGCAACTGCTA 70
128 ATGGGCGAGATTAACTGCAAGACCACTGCAACGGCACTCTGTTGAACGCCGCA 187
71 ATGGGCGAGATTAACTGCAAGACCACTGCAACGGCACTCTGTTGAACGCCGCA 130
188 TCACCTTACATATAGGAAGGCGCTCCATTAATTAATGACCAATGGAACATGATA 247
131 TCACCTTACATATAGGAAGGCGCTCCATTAATTAATGACCAATGGAACATGATA 120
248 CAGATTAATCTTGGTGGACCTTACTCTTGGCGCTTAATAACAGAACACATGACAT 307
191 CAGATTAATCTTGGTGGACCTTACTCTTGGCGCTTAATAACAGAACACATGACAT 250
308 CTGTAATCTTGAAGATGCGGCGGTGATCCATTAATTAATTAATTAATTAATTAAT 367
251 CTGTAATCTTGAAGATGCGGCGGTGATCCATTAATTAATTAATTAATTAATTAAT 310
368 GAAATTAACCAACCAATTCATCAAAAGTATGGAAGATTTAGAGATTTGATATCT 427
311 GAAATTAACCAACCAATTCATCAAAAGTATGGAAGATTTAGAGATTTGATATCT 370
428 CTGTAATCTTGAAGATGCGGCGGTGATCCATTAATTAATTAATTAATTAATTAAT 487
371 CTGTAATCTTGAAGATGCGGCGGTGATCCATTAATTAATTAATTAATTAATTAAT 430
488 AGGATTTCTTGGTGGACCTTACTCTTGGCGCTTAATAACAGAACACATGACAT 547
431 AGGATTTCTTGGTGGACCTTACTCTTGGCGCTTAATAACAGAACACATGACAT 490
548 CTGATGCTTAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAAC 598
491 CTGATGCTTAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAAC 550
599 ACGAACAACCAACCAATTTATGAGATCTAGAAAAGTATGAGGATGGAACAATTA 658
551 ACGAACAACCAACCAATTTATGAGATCTAGAAAAGTATGAGGATGGAACAATTA 610
659 CTGATGCTTAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAAC 718
611 CTGATGCTTAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAAC 670
719 ATATTCATTAATATATGATCCGAAAGATTCGAGGATTTTGGCAATAGAGATTA 778
671 ATATTCATTAATATATGATCCGAAAGATTCGAGGATTTTGGCAATAGAGATTA 730
779 CTGATGCTTAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAAC 838
731 CTGATGCTTAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAACCGTAAAC 790
839 TTAAGGCGGTGATTAAGAACCAACCAATGCAAAATGCAACGCAAGGATTAAGG 898
791 TTAAGGCGGTGATTAAGAACCAACCAATGCAAAATGCAACGCAAGGATTAAGG 850
899 TTGATCATGCGCGAATACCAATGCAAAATGCAACGCAAGGATTAAGGATTAAG 958
851 TTGATCATGCGCGAATACCAATGCAAAATGCAACGCAAGGATTAAGGATTAAG 910
959 CTTTAACCGCGATCGATGATGATGATATTTTGAAGATGATGATGATGATGATG 1018
911 CTTTAACCGCGATCGATGATGATGATATTTTGAAGATGATGATGATGATGATG 970
1019 ATCATGCTGATGCGCGAATGCAAAATGCAACGCAAGGATTAAGGATTAAGG 1078
971 ATCATGCTGATGCGCGAATGCAAAATGCAACGCAAGGATTAAGGATTAAGG 1030
1079 AGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
1031 AGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1090
QY 1139 ATATCATGATTAATCATTAATCATGATGATGATGATGATGATGATGATGATGAT 1198
|||||

```

	Db	1091	ATATCATGATTAATTCATTATCTCAGTAGATAAAAAATTATATGTGTATACAGATGAAG	1150
Oy	1199	AGATTTTTAAAGCCAGGCCAGTGTCAAAATTCTCGAGAAAATGCTCAAACACCATTTC	1258	
Db	1151	AGATTTTTAAAGCCAGGCCAGTGTCAAAATTCTCGAGAAAATGCTCAAACACCATTTC	1210	
Oy	1259	AGGTGATGGAGCAGAAAAAATATGCTGCATATAAGACTTCTTGGAGGTGAAATA	1318	
Db	1211	AGGTGATGGAGCAGAAAAAATATGCTGCATATAAGACTTCTTGGAGGTGAAATA	1270	
Oy	1319	AAAGAACCAGGGCGTGGGTATCTGAAAAAACTCATCTTATATATAAGATAAATG	1378	
Db	1271	AAAGAACCAGGGCGTGGGTATCTGAAAAAACTCATCTTATATATAAGATAAATG	1330	
Oy	1379	AATTTCTCACCTTCCCACACACCTCTCGGATTAATCCAGGCCGCTCAACGACTTGTTA	1438	
Db	1331	AATTTCTCACCTTCCCACACACCTCTCGGATTAATCCAGGCCGCTCAACGACTTGTTA	1390	
Oy	1439	GCCAAAAACAACACTGAGCTGTCTGATTTATCATCAGTTTTTATTCACGATTGAAGCAC	1498	
Db	1391	GCCAAAAACAACACTGAGCTGTCTGATTTATCATCAGTTTTTATTCACGATTGAAGCAC	1450	
Oy	1499	TGACCGTTTCATTCAGAAATATGATTCAGTGAATGCAACGCTGCTAGATGACAGCTGTG	1558	
Db	1451	TGACCGTTTCATTCAGAAATATGATTCAGTGAATGCAACGCTGCTAGATGACAGCTGTG	1510	
Oy	1559	GTAATGTA 1566		
Db	1511	GTAATGTA 1518		
RESULT	2			
LOCUS	A56808	1462 bp	DNA	PAT 03-MAR-1998
DEFINITION	Sequence 16 from Patent WO9628551.			
ACCESSION	A56808			
VERSION	A56808.1	GI:3712821		
KEYWORDS				
SOURCE	Yersinia pestis.			
ORGANISM	Yersinia pestis.			
REFERENCE	Bacteria: Proteobacteriae; gamma subdivision; Enterobacteriaceae;			
AUTHORS	Yesinia.			
TITLE	1 (bases 1 to 1462)			
JOURNAL	Tilball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,			
COMMENT	Bennett and Alice,M.			
FEATURES	VACCINES FOR PLAGUE			
SOURCE	Patient: NO 9628551-A 16 19-SEP-1996;			
	SECR DEFENCE (GB)			
	Other publication ZA 9602036 960716			
	Other publication AU 4951196 961002.			
	Location/Qualifiers			
	1..1462			
	/organism="Yersinia pestis"			
	/db_xref="taxon:632"			
	<8..1450			
	/note="unnamed protein product"			
	/codon_start=1			
	/transl_table=1			
	/protein_id="CAA03421.1"			
	/db_xref="gi:3712822"			
	/translation="ADLTASTATATATLVEPARITLYKEGAPTIMNGNIDTELLG			
	TUTIGKRTGTSISVNFETDAGDPMTILFTFSOGNNHOFITTKIGDSRDFDISPRV			
	NGENVLGDVDVALTGSDQFVFRTSGSGGLAAQRKYDAIVAYVSNGSLGRRAVE			
	ONPHFELEDEKRVEDLTGHSSVSELVOLVDKNIDISIKYDKPDRSDSFVANRYT			
	IDDLLELKLIAYFLPEDAILKGHYDNONGIKRYAEFLSPNTOMELRAPAMAVWY			
	HFSFLAIRDDIDILKIYVDSNMHHGDARSKLRELAELTALFKIYVIOEINKHLS			
	CSLTINHDSKTINIMDKNIYGYTDPEIRKASLYVTLKMPOTTYOVGSESKYISID			
	FISEENKRTGALGNLKNYSYRNKRNNELSHPAITCDKSRLPLNDLVASQTITQJSDITS			
	RNSAIETLNLFIEKRIYDSVMORLLDDDTSGK"			
BASE COUNT	476 a	286 c	300 g	400 t
ORIGIN				

Query Match	Best Local Similarity	90.1%	Score 1410.6	DB 5	Length 1462
Matches 1432	Conservative	0	Mismatches 295	Indels	Gaps
Qy	131	CGGAGAGTTTAAACGCAACGACACCTGCAACGCGCACTCTTGTGACACGCGCCGATCA	190		
Db	6	CGGAGAGTTTAAACGCAACGACACCTGCAACGCGCACTCTTGTGACACGCGCCGATCA	65		
Qy	191	CTCTTACATATTAAGGAAGGCGCTCCAAATTACATTTATGACAATGGAACATCGATACAG	250		
Db	66	CTCTTACATATTAAGGAAGGCGCTCCAAATTACATTTATGACAATGGAACATCGATACAG	125		
Qy	251	AATTAAGTGTGGAGCGCTTACCTCTGCGGCGCTATATAACAGGAACCACTAGCAATCTG	310		
Db	126	AATTAAGTGTGGAGCGCTTACCTCTGCGGCGCTATATAACAGGAACCACTAGCAATCTG	185		
Qy	311	TTAACTTTACAGATGCGCGGCGGTGATCCCATGACTTAACTTACTCTCGATGGAA	370		
Db	186	TTAACTTTACAGATGCGCGGCGGTGATCCCATGACTTAACTTACTCTCGATGGAA	245		
Qy	371	ATAACACCAATTCACCTACCAAAAGTATTGGCAGAGATTCTAGAGATTTTGATCTCTC	430		
Db	246	ATAACACCAATTCACCTACCAAAAGTATTGGCAGAGATTCTAGAGATTTTGATCTCTC	305		
Qy	431	CTAAGGTAAACGGTGAAGAACTTGTGGGGGATGACCTGCTTGGCTACGGCACCCAG	490		
Db	306	CTAAGGTAAACGGTGAAGAACTTGTGGGGGATGACCTGCTTGGCTACGGCACCCAG	365		
Qy	491	ATTCTCTTGTGCGCTCAATTGTTCCAAAGGCGGTAAATTGSCACAGGTAAATACACTG	550		
Db	366	ATTCTCTTGTGCGCTCAATTGTTCCAAAGGCGGTAAATTGSCACAGGTAAATACACTG	425		
Qy	551	ATGCTGTAAACCGTAACCGTATCTAAACCAAGATTCAAT-----GATTAGCGCTACG	601		
Db	426	ATGCTGTAAACCGTAACCGTATCTAAACCAAGATTCAATCGAAGTGGTATTAGAGCTACG	485		
Qy	602	AACAAACCCCAACAACTTTATTGAGAGTCTAGAAAABTTAGGTTGGAACAACCTACTG	661		
Db	486	AACAAACCCCAACAACTTTATTGAGAGTCTAGAAAABTTAGGTTGGAACAACCTACTG	545		
Qy	662	GTCATGTTCTCAGTTTATGAGAATTTGGTTCAGTTAGTCAAAAGTATAAAATATATAGTA	721		
Db	546	GTCATGTTCTCAGTTTATGAGAATTTGGTTCAGTTAGTCAAAAGTATAAAATATATAGTA	605		
Qy	722	TTTCCATTAATATGATCCAGAAAAGATTTCGAGGTTTTCGCCAATAGAGTAATTAAGT	781		
Db	606	TTTCCATTAATATGATCCAGAAAAGATTTCGAGGTTTTCGCCAATAGAGTAATTAAGT	665		
Qy	782	ATGATATCGAATTCCTAGAAAATTCCTAGCTTATTTTACCAGAGATACCAATCTTA	841		
Db	666	ATGATATCGAATTCCTAGAAAATTCCTAGCTTATTTTACCAGAGATACCAATCTTA	725		
Qy	842	AAGCGGCTATATGACAACCACTCAAAAATGAGGATCAAGGAGTAAAGATGCTCTG	901		
Db	726	AAGCGGCTATATGACAACCACTCAAAAATGAGGATCAAGGAGTAAAGATGCTCTG	785		
Qy	902	AATCATGCGCGAATACCAATGGAATTTGCGGCGCTTACGACAGTAATGCAATTTCTTT	961		
Db	786	AATCATGCGCGAATACCAATGGAATTTGCGGCGCTTACGACAGTAATGCAATTTCTTT	845		
Qy	962	TAACCGCGGATCGATGATGATGATATTTTGAAAGTATGTTGATTCATGATATC	1021		
Db	846	TAACCGCGGATCGATGATGATGATATTTTGAAAGTATGTTGATTCATGATATC	905		
Qy	1022	ATGAGTATGCGCGAGCAATGTCGGAAGAAATTTAGCTAGTTCACCGCGAATTAAGA	1081		
Db	906	ATGAGTATGCGCGAGCAATGTCGGAAGAAATTTAGCTAGTTCACCGCGAATTAAGA	965		
Qy	1082	TTTATTAAGTATATCAAGCCGAATTAATTAAGATCTGCTAGTATGCGACCATTAATA	1141		
Db	966	TTTATTAAGTATATCAAGCCGAATTAATTAAGATCTGCTAGTATGCGACCATTAATA	1025		
Qy	1142	TCGATGATTAATCCATTAATCTCATGAGATAAAAATTTATGCTTATACATGATAGAGA	1201		

```

Db 1026 TCCATATTAATCCATTAACTATGATGATTAATAATTTATATGTTATACAGATGAAGAGA 1085
QY 1202 TTTTAAAGCCAGCGAGAGTACAAATCTCTCGAGAAAATGCTCAACCAACCATTCAGG 1261
Db 1086 TTTTAAAGCCAGCGAGAGTACAAATCTCTCGAGAAAATGCTCAACCAACCATTCAGG 1145
QY 1262 TGGATGGAGCGAGAAAATAGTCTCGATAAAGACATTTCTTGGAAAGTGAAGATAAAA 1321
Db 1146 TGGATGGAGCGAGAAAATAGTCTCGATAAAGACATTTCTTGGAAAGTGAAGATAAAA 1205
QY 1322 GAACCGGGGCGTGGTAACTGTAATAAATCAATCTTATATAAAGATAAATGAAT 1381
Db 1206 GAACCGGGGCGTGGTAACTGTAATAAATCAATCTTATATAAAGATAAATGAAT 1265
QY 1382 TATCTCACTTTGGCCACCACTGCTCGGATTAATCAAGCGCGCTCAACGACTTGGTAGCC 1441
Db 1266 TATCTCACTTTGGCCACCACTGCTCGGATTAATCAAGCGCGCTCAACGACTTGGTAGCC 1325
QY 1442 AAAAAACAACCTGCTGCTGATATTAATCAATCAAGCTTTTAAATGAGTATGAAGACTGA 1501
Db 1326 AAAAAACAACCTGCTGCTGATATTAATCAATCAAGCTTTTAAATGAGTATGAAGACTGA 1385
QY 1502 ACCGTTTCAATGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1561
Db 1386 ACCGTTTCAATGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1445
QY 1562 AATGA 1566
Db 1446 AATGA 1450

RESULT 3
LOCUS YEPICR 2100 bp DNA BCT 26-APR-1993
DEFINITION Yersinia pestis lcrG, lcrV, and lcrH genes, complete cds.
ACCESSION M26405
VERSION M26405.1 GI:155448
KEYWORDS lcrG protein; lcrH protein; lcrV protein; V antigen.
SOURCE 75kb virulence plasmid.
ORGANISM Yersinia pestis
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 2100)
AUTHORS Price, S.B., Leung, K.Y., Bayre, S.S. and Straley, S.C.
TITLE Molecular analysis of lcrGVH, the V antigen operon of Yersinia
Pestis
JOURNAL J. Bacteriol. 171, 5646-5653 (1989)
MEDLINE 9008806
COMMENT Computer readable copy of sequence [J. Bacteriol. (1989) In press]
kindly submitted by Price, S.
07-AUG-1989.

FEATURES
source location/Qualifiers
1..2100
/organism="Yersinia pestis"
/strain="KIMS"
/db_xref="taxon:632"
-35_signal 147..152
-10_signal 174..179
gene 192..479
CDS 192..479
/gene="lcrG"
/gene="lcrH"
/codon_start=1
/transl_table=11
/protein_id="AAA27640.1"
/db_xref="GI:155449"
/translation="MKSSHFDEYDKTLKQAEIAIADSDHRATLQEMCADILPEAV
MKIFGRSAEEIKPAERELIDIKRQEROPQHPYDKRPRKPTMMRGII"
RBS 463..466
/gene="lcrG"
/gene="lcrH"
463..1461
/gene="lcrV"

```

```

CDS
481..1461
/gene="lcrV"
/codon_start=1
/transl_table=11
/protein_id="AAA27641.1"
/db_xref="GI:155450"
/translation="WIRAYEONPOHFIEDLEKRVYOLTGHSVLEELVOLVMDKNT
DISIKYDPRKSEVFANVITDDIELKRIILAYLPEDAILKGGHYDQNLQGIKRV
EFLESPNTOMELRAFMAVMEHSLIADRIDDIKLYVDSNMHGHGASKIREELAE
TAEIKYIVIAOEIKHLSGGTINIHDKSINIMDKNLGYTDEIEFASAEYKLEK
MPOITIOVDSSEKRVISIKDPLSEKRGALGNKNSYVKNPNEISHRATICSOK
SRPLNDVSOQTQLSDITSRPNSAIEALNRFIOKYVDSVMORLLDDISGK"
RBS
1465..1468
/gene="lcrH"
1465..1980
/gene="lcrH"
1474..1980
CDS
/gene="lcrH"
/codon_start=1
/transl_table=11
/protein_id="AAA27642.1"
/db_xref="GI:155451"
/translation="MQQETPTQOELAMESFLKGGTIANLEISSTLEQVSLAF
NOYSGKTEDHAKYFQALCVLDHDSRFELIGACRQMGQYDIAIHSTYGALMDIK
EPRFPHAECLLOKGELEAESGFLIADIELADTERKELSTRVSMLEAIKIKREM
EHECVNDP"
BASE COUNT 688 a 423 c 461 g 528 t
ORIGIN

Query Match 62.5%; Score 979.4; DB 1; Length 2100;
Best Local Similarity 99.9%; Pred. No. 4.7e-202;
Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 586 ATGATTAGAGCCTAGCAAAAAAACCACAACTTTTATGAGATCTACAAAAAGTTAGG 645
Db 481 ATGATTAGAGCCTAGCAAAAAAACCACAACTTTTATGAGATCTACAAAAAGTTAGG 540
QY 646 GGGCAACAATCTACTGCTGATGCTTCTTCAAGTTTGAAGAAATGGTTAGTAGCAAA 705
Db 541 GGGCAACAATCTACTGCTGATGCTTCTTCAAGTTTGAAGAAATGGTTAGTAGCAAA 600
QY 706 GATTAATAATATGATATTTCCATTAATATGATCCGAGAAAGATTCGAGGTTTGGCC 765
Db 601 GATTAATAATATGATATTTCCATTAATATGATCCGAGAAAGATTCGAGGTTTGGCC 660
QY 766 AATAGGTAATACGTATATATCGAATTCGATGCTCAAGAAATCCTAGCTTATTTTACCC 825
Db 661 AATAGGTAATACGTATATATCGAATTCGATGCTCAAGAAATCCTAGCTTATTTTACCC 720
QY 826 GAGATACCATCTTTAAAGCGGCTATATGACAAACCACTGCAAAATGGCATCAAGCA 885
Db 721 GAGATACCATCTTTAAAGCGGCTATATGACAAACCACTGCAAAATGGCATCAAGCA 780
QY 886 GAAAAAGAGTCTTGAATCATCGCGGAATACACAATGGGAATGGCGGCTTCATGGCA 945
Db 781 GAAAAAGAGTCTTGAATCATCGCGGAATACACAATGGGAATGGCGGCTTCATGGCA 840
QY 946 GTAATGCAATTTCTTTAAACCGCGGATCGATGATGATGATGATGATGATGATGATGAT 1005
Db 841 GTAATGCAATTTCTTTAAACCGCGGATCGATGATGATGATGATGATGATGATGATGAT 900
QY 1006 GATTCAGTAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
Db 901 GATTCAGTAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 1066 ACCGCGGAATTAAGATTATTCAGTTATTCAGCGGAATTAATTAAGCATCTGCTAGT 1125
Db 961 ACCGCGGAATTAAGATTATTCAGTTATTCAGCGGAATTAATTAAGCATCTGCTAGT 1020
QY 1126 AGTGGCACCATAAATATCCATGATTAATTCATGATGATGATGATGATGATGATGATGAT 1185
Db 1021 AGTGGCACCATAAATATCCATGATTAATTCATGATGATGATGATGATGATGATGATGAT 1080

```

QY 1186 TATACAGATGAGAGATTTTTAAAGCCAGCCAGAGTACAAATTCGAGAAATGCT 1245
 |||||||
 Db 1081 TATACAGATGAGAGATTTTTAAAGCCAGCCAGAGTACAAATTCGAGAAATGCT 1140
 |||||||
 QY 1246 CAACACCATTCAGAGTGGATGGAGGAGAGAAAAAATAGTCTCGATAAGACTTTCT 1305
 |||||||
 Db 1141 CAACACCATTCAGAGTGGATGGAGGAGAGAAAAAATAGTCTCGATAAGACTTTCT 1200
 |||||||
 QY 1306 GGAATGAGATTAAGAAACCGGGCGCTGGTATCTGAAAAATCTACTCTTATAT 1365
 |||||||
 Db 1201 GGAATGAGATTAAGAAACCGGGCGCTGGTATCTGAAAAATCTACTCTTATAT 1260
 |||||||
 QY 1366 AAGATTAATGATATATCTCATCTTCCACACCTCTCGATAATCCAGCGGCTC 1425
 |||||||
 Db 1261 AAGATTAATGATATATCTCATCTTCCACACCTCTCGATAATCCAGCGGCTC 1320
 |||||||
 QY 1426 AACGACTTGGTTAGCCAAAAAACAACCTGCTGTGATATATACATCAGTTTATTC 1485
 |||||||
 Db 1321 AACGACTTGGTTAGCCAAAAAACAACCTGCTGTGATATATACATCAGTTTATTC 1380
 |||||||
 QY 1486 GCTATGAGACACGACCCGTTTCTATGAGAAATGATTCAGTATGCAAGCTGCTA 1545
 |||||||
 Db 1381 GCTATGAGACACGACCCGTTTCTATGAGAAATGATTCAGTATGCAAGCTGCTA 1440
 |||||||
 QY 1546 GATGACACGCTGTGTAATGA 1566
 |||||||
 Db 1441 GATGACACGCTGTGTAATGA 1461
 |||||||

RESULT 4

YPCD1/c 70305 bp DNA BCT 22-MAR-2000
 DEFINITION Yersinia pestis plasmid pCD1.
 VERSION AL117189.1 GI:5832423
 KEYWORDS chapomex; cytotoxic effector; IS100; IS1616; IS1617; lcr;
 low-calcium response; ysc; targeted effector; transposase; type III
 secretion; Y antigen; virulence; ylp; yop; ysc.

SOURCE

ORGANISM Yersinia pestis
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.

REFERENCE 1 (bases 1 to 70305)
 AUTHORS Karlyshev, A.V. and Wren, B.W.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 70305)
 AUTHORS Baker, S.G. and Mungall, K.
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 70305)
 AUTHORS James, K.D., Parkhill, J., Barrell, B.G. and Rajandream, M.A.
 JOURNAL Direct Submission
 TITLE Submitted (03-SEP-1999) Yersinia pestis sequencing project, Sanger
 Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA
 E-mail: barrell@sanger.ac.uk DNA supplied by Dr. Andrey Karlyshev
 and Prof. Brendan Wren. [3]. Department of Infectious and Tropical
 Diseases, London School of Hygiene and Tropical Medicine, Keppel
 Street, London WC1E 7HT

COMMENT

Notes:
 Yersinia pestis sequencing at The Sanger Centre is funded by
 Beowulf Genomics
 Details of Y. pestis sequencing at the Sanger Centre are available
 on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/Y_pestis/)
 CDS are numbered using the following system eg YPCP1.01c. YP (Y.
 pestis), PCP1 (Plasmid name), .01 (first CDS), c (complementary
 strand).
 The more significant matches with motifs in the PROSITE database
 are also included but some of these may be fortuitous.
 The length in codons is given for each CDS.
 Usually the highest scoring match found by fasta -o is given for
 CDS which show significant similarity to other CDS in the database.
 The position of possible ribosome binding site sequences are given

where these have been used to deduce the initiation codon.
 CAUTION: We may not have predicted the correct initiation codon.
 Where possible we choose an initiation codon (atg, atg, ttg or
 att) which is preceded by an upstream ribosome binding site
 sequence (optimally 5-13bp before the initiation codon). If this
 cannot be identified we choose the most upstream initiation codon.
 Location/Qualifiers
 1. 70305

FEATURES

source

repeat_unit

gene

CDS

misc_feature

gene

CDS

misc_feature

/organism="Yersinia pestis"
 /plasmid="pCD1"
 /strain="CO-92 Blover Orientalis"
 /db_xref="taxon:632"
 1. 1996
 /note="IS100 element"
 88. 1110
 /gene="YPCD1.01"
 88. 1110
 /gene="YPCD1.01"
 /note="YPCD1.01, probable transposase, len: 340 aa;
 putative insertion sequence IS100, identical to
 corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946,
 AF074612) (340 aa), fasta scores: opt: 2328 z-score:
 3808.9 E(): 0, 100.0% identity in 340 aa overlap. Similar
 to many others e.g. TRA0_ECOLI (EMBL:X14793), lcr, E.coli
 transposase for insertion sequence element IS21 (390 aa)
 (33.1% identity in 329 aa overlap). Contains Pfam match to
 entry PF00239 recombinase, site-specific recombinases,
 score 25.70, E-value 4.8e-06. Contains probable
 helix-turn-helix motif at aa 19-40 (Score 2045, +6.15 SD)."
 /codon_start=1
 /transl_table=11
 /label="YPCD1.01"
 /product="putative transposase"
 /protein_id="CAB54878.1"
 /db_xref="GI:5832424"
 /translation="MVFETVMEIKILKMGSSRAARELGISRNTKRYLQAKSEP
 PKTPRAVAVSLDEVDYRORADHPKIPATVAREIRGGYGGGATIRAFR
 SLSPQROEAVAFVETPEGRMOVDNCTMNGSPILFPAVAVGYSMLYIEFDNR
 YDLETRHRAAFRGVGVREYLDNKKTYVLORDATQOCHRPHPRLPGFKMGFS
 PRCPFRAGTKGVKERVQYTRNSFTPLMTLRPGITGVETANRHLRLHADV
 NORKHETIQARPCDRLMEQSMALPPEKKEYDVHIDENLVNFKRHLHPSTYDS
 FCRGVA"
 112. 195
 /gene="YPCD1.01"
 /note="Pfam match to entry PF00239 recombinase,
 site-specific recombinases, score 25.70, E-value 4.8e-06"
 1110. 1889
 /gene="YPCD1.02"
 1110. 1889
 /gene="YPCD1.02"
 /note="YPCD1.02, probable transposase, len: 259 aa;
 putative insertion sequence IS100, identical to
 corresponding CDS from Y.pestis KIM5 pCD1 (EMBL:AF053946,
 AF074612) (259 aa), fasta scores: opt: 1658 z-score:
 2693.4 E(): 0, 100.0% identity in 259 aa overlap. Similar
 to many others e.g. lcr_ECOLI (EMBL:X14793), lcr, E.coli
 transposase for insertion sequence element IS21 (265 aa)
 (47.4% identity in 249 aa overlap). Contains PF00017
 ATP/GTP-binding site motif A (P-loop)."
 /codon_start=1
 /transl_table=11
 /label="YPCD1.02"
 /product="putative ATP-binding protein"
 /protein_id="CAB54879.1"
 /db_xref="GI:5832425"
 /translation="MNEIQRHMLALAGOLESLISAPALSOAVDQSVYDFLE
 HILHEKRLARHQRQAMVTRMAFPATKEEYDFDPAPOKQSLSTIEFN
 ENIVLSPSVGRKTHLAINGYAVRAGIVRTTAADLLQSTRAQRQRYKTLLR
 GVAVPRLIIDEIGYLFSEDEAKLFFQVIAKREKSAMLITNLPFGQWDQTFAGA
 ALTSAMDRLIHSHVQIKGESYRLNOKRKAGVIAENPE"
 1434. 1457
 /gene="YPCD1.02"
 /note="P00017 ATP/GTP-binding site motif A (P-loop)."

[illegible]

[illegible]

```

/codon_start=-1
/transl_table=-1
/product="Yop proteins translocation protein I homolog"
/protin_id="A062608.1"
/db_xref="GI:2996285"
/translation="MPRIEADQDEVITITLEELGPAEPTDQIMKPDAMSDTQCL
GSHLEVSIDQISFKTVKSDHLTKLAVSDNDMLMOWMSLIRITIOELLAKTAG
RMSQNVETLSKGS"
complement(6094. .6591)
/gene="yscH"
complement(6094. .6591)
/gene="yscH"
/codon_start=-1
/transl_table=-1
/product="Yop proteins translocation protein H homolog"
/protin_id="A062547.1"
/db_xref="GI:2996224"
/translation="MTVTTLNRGSITSLMSSQAVSTLQPVASELKTQLENKLSSEAE
KREVLMOQIYASNPDPHAYLEVATPVRALLARRGQGVPAIDPELRSVLAQ
FDSFGKRWEMAILLQVLEGIKPNESQVGLPYLSLILINKELMILLPNSIYDSLNSHQ
IDMDT"
complement(6588. .6935)
/gene="yscG"
complement(6588. .6935)
/gene="yscG"
/codon_start=-1
/transl_table=-1
/product="Yop proteins translocation protein G homolog"
/protin_id="A062548.1"
/db_xref="GI:2996225"
/translation="MKIKLVLAETALIGTGHYHEANCIAEMHLKGEPAVOLI
RUSLSNRGDYASALQGNKRLAYPDLEPWLALCEYRLGLSALLESIRLARSQDRI
OTFVNMREQLKT"
complement(6937. .7200)
/gene="yscF"
complement(6937. .7200)
/gene="yscF"
/codon_start=-1
/transl_table=-1
/product="Yop proteins translocation protein F homolog"
/protin_id="A062549.1"
/db_xref="GI:2996226"
/translation="MSNFSGFTKGTIDADIADAQVLKRPADANKAVNDSIALDK
PNDPALADQHSINMSVYINISIVSMKLMOGIIOKFP"
complement(7201. .7401)
/gene="yscE"
complement(7201. .7401)
/gene="yscE"
/codon_start=-1
/transl_table=-1
/product="Yop proteins translocation protein E homolog"
/protin_id="A062550.1"
/db_xref="GI:2996227"
/translation="MTQLEBQLHNVETVASTITQLEMALTKLKKDMKRGDAKQYQW
QRESKALSAIAIHTVAGDK"
complement(7398. .8657)
/gene="yscD"
complement(7398. .8657)
/gene="yscD"
/codon_start=-1
/transl_table=-1
/product="Yop proteins translocation protein D homolog"
/protin_id="A062551.1"
/db_xref="GI:2996228"
/translation="MSWCBRYFGKRGVVEELPHGRVYGSPIQSDIYLSNSETAP
VHLVAVDEGRIETLSAEPLDEGLPYPLGTLIRAGSCIEVFLMTYVAVGQPE
TLDVPTQKEPTDRLPRSRIGLSVLSLILTLGLGHLGRLRYNDQGVLEQEV
RLRLAAYKADVLTSPKKEGEPWLTGYIQDNHARLSLONFLESIGIFRLRAME
ELRQAGETLQRLRYGIEVSLAPQAGMQLONNEVESEIQKIDSLDLAEEVGLGLG
ESKVRILAGNQRRKRLDALBEQFGLSDFTYVAVKELLELRQGVNDEKLNSNQQQYFR

```

Query Match 62.58; Score 979.4; DB 2; Length 70504;
 Best Local Similarity 99.98; Pred. No. 4e-202;

Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 586 ATGATTAGACCTACGACAAACCCACACATTTATTGAGATCTAGAAAAAGTTAG 645
|||||
Db 25311 ATGATTAGACCTACGACAAACCCACACATTTATTGAGATCTAGAAAAAGTTAG 25370
OY 646 GTGACACACTACTGCTCATGTTCTTCAGTTTGAAGAATGGTTCAGTTAGTCA 705
|||||
Db 25371 GTGACACACTACTGCTCATGTTCTTCAGTTTGAAGAATGGTTCAGTTAGTCA 25430
OY 706 GATATAATATATATATTTCCATTAAATATGATCCAGAAAATTCGGAGTTTGGC 765
|||||
Db 25431 GATATAATATATATATTTCCATTAAATATGATCCAGAAAATTCGGAGTTTGGC 25490
OY 766 AATAGATAATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 825
|||||
Db 25491 AATAGATAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25550
OY 826 GAGATACCATCTTAAAGCGGTCATATGACACCACTGCAAAATGGCATCAAGCG 885
|||||
Db 25551 GAGATACCATCTTAAAGCGGTCATATGACACCACTGCAAAATGGCATCAAGCG 25610
OY 886 GTAAAAAGTTCCTGATATATATGATGATGATGATGATGATGATGATGATGAT 945
|||||
Db 25611 GTAAAAAGTTCCTGATATATATGATGATGATGATGATGATGATGATGATGAT 25670
OY 946 GTATGATTTCTCTTTAACCGCGATCGATGATGATGATGATGATGATGATGATGAT 1005
|||||
Db 25671 GTATGATTTCTCTTTAACCGCGATCGATGATGATGATGATGATGATGATGATGAT 25730
OY 1006 GATTCATGATATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
|||||
Db 25731 GATTCATGATATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 25790
OY 1066 ACCGCCAATTAAGATTTATTCAGTATTCAGCCGAATTAAGATCTGCTAGT 1125
|||||
Db 25791 ACCGCCAATTAAGATTTATTCAGTATTCAGCCGAATTAAGATCTGCTAGT 25850
OY 1126 AGTGGCACATTAATATTCATGATTAATCATTAATCTCATGATTAATTAATG 1185
|||||
Db 25851 AGTGGCACATTAATATTCATGATTAATCATTAATCTCATGATTAATTAATG 25910
OY 1186 TATACAGATGAAGATTTTAAAGCCAGGAGATACAAATTCGAGAAATGCGT 1245
|||||
Db 25911 TATACAGATGAAGATTTTAAAGCCAGGAGATACAAATTCGAGAAATGCGT 25970
OY 1246 CAACACCATTCAGTGGATGGAGCGAGAAAAAATAGTCTCGATTAAGACTTTCT 1305
|||||
Db 25971 CAACACCATTCAGTGGATGGAGCGAGAAAAAATAGTCTCGATTAAGACTTTCT 26030
OY 1306 GGAAGTGAATTAAGAACACGGGGCGTGGTAACTGTGAATAAATCATCTTAAAT 1365
|||||
Db 26031 GGAAGTGAATTAAGAACACGGGGCGTGGTAACTGTGAATAAATCATCTTAAAT 26090
OY 1366 AAGATTAATTAATATATCTCATCTTGCACACCTGCTCGATAAGTCCAGCCGCT 1425
|||||
Db 26091 AAGATTAATTAATATATCTCATCTTGCACACCTGCTCGATAAGTCCAGCCGCT 26150
OY 1426 AAGACTTGTGTTAGCCAAAAACAACACTGCTGTGATATTAACACGTTTAAAT 1485
|||||
Db 26151 AAGACTTGTGTTAGCCAAAAACAACACTGCTGTGATATTAACACGTTTAAAT 26210
OY 1486 GCTATTGAACACTGAACCGTTTCATTCAGAAATATGATCACTGATGCAAGCTGCT 1545
|||||
Db 26211 GCTATTGAACACTGAACCGTTTCATTCAGAAATATGATCACTGATGCAAGCTGCT 26270
OY 1546 GATGACACGCTGCTGTAATGA 1566
|||||
Db 26271 GATGACACGCTGCTGTAATGA 26291

```

RESULT 6
 AF074612

LOCUS AF074612 70559 bp DNA circular BCT 07-APR-2000
 DEFINITION Yersinia pestis plasmid pCD1, complete plasmid sequence.
 ACCESSION AF074612 M25810
 VERSION AF074612.1 GI:3822037
 KEYWORDS Yersinia pestis.
 ORGANISM Yersinia pestis.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.
 REFERENCE 1 (bases 4338 to 44840)
 AUTHORS Leung,K.Y. and Straley,S.C.
 TITLE The yopM gene of Yersinia pestis encodes a released protein having homology with the human platelet surface protein GPIIb alpha J. Bacteriol. 171 (9), 4623-4632 (1989)
 JOURNAL MEDLINE 89359090
 REFERENCE 2 (bases 1 to 70559)
 AUTHORS Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and Blattner,F.R.
 TITLE DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia pestis KIM5 Infect. Immun. 66 (10), 4611-4623 (1998)
 JOURNAL MEDLINE 98427122
 REFERENCE 3 (bases 4338 to 44840)
 AUTHORS Straley,S.C.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-1993) Microbiology and Immunology, University of Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA
 REFERENCE 4 (bases 1 to 70559)
 AUTHORS Perry,R.D., Straley,S.C., Fetherston,J.D., Rose,D.J., Gregor,J. and Blattner,F.R.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUN-1998) Microbiology and Immunology, University of Kentucky, M5415 Medical Center, Lexington, KY 40536-0084, USA
 COMMENT On Apr 7, 2000 this sequence version replaced gi:155539.
 FEATURES
 source location/Qualifiers
 1. .70559
 /organism="Yersinia pestis"
 /plasmid="pCD1"
 /strain="KIM5"
 /db_xref="taxon:632"
 57. .368
 /gene="Y0001"
 57. .368
 /note="Y0103: 43 pct identical (0 gaps) to 100 residues of an approx. 200 aa protein GENEPEPT: g1|537126, ori_0198 Escherichia coli"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAC69758.1"
 /db_xref="GI:3822038"
 /translation="MHQOSRGAAGSRTLSILMROSGINVVRLARLRRCGLASROP GKPRYGEREVSILASPDLLKRFKPSPEPNVWSGYISYIKVNGWCYLAIVLYSFH W"
 655. .1033
 /gene="nuc"
 /note="Y0002"
 655. .1033
 /gene="nuc"
 /codon_start=1
 /transl_table=11
 /product="endonuclease"
 /protein_id="AAC69759.1"
 /db_xref="GI:3822039"
 /translation="NDIKLQHTPIGTMDYRPVNRKSGGRLRRCPPDFVIHYRMDLV VNGIGFVATVNSFKALHDKYIIVDKNTQWGSFNSQAQVQSENVLITMGDFTVQV AYLOYOSRWNRKGTDMRSSY"
 1171. .1425
 /gene="repB"
 /note="Y0003"
 1171. .1425
 /gene="repB"

misc_RNA
 /codon_start=1
 /transl_table=11
 /product="replication protein"
 /protein_id="AAC69760.1"
 /db_xref="GI:3822040"
 /translation="MSQIENAVTSSSRKRAYRGNPLTGAERKQMSYRRKRETHKAINV FIQNDLKNELDCEDSGLTQTEMIRWIOREKARTNAA"
 complement(1560, 1649)
 /note="antisense RNA"
 /product="copA"
 1667. .1741
 /gene="tap"
 /note="Y0004"
 1667. .1741
 /gene="tap"
 /codon_start=1
 /transl_table=11
 /product="repa translation protein"
 /protein_id="AAC69761.1"
 /db_xref="GI:3822041"
 /translation="MKRKQVLLRLLLLPENISAGKCD"
 1734. .2600
 /gene="repa"
 /note="Y0005"
 1734. .2600
 /gene="repa"
 /codon_start=1
 /transl_table=11
 /product="replication protein"
 /protein_id="AAC69762.1"
 /db_xref="GI:3822042"
 /translation="MTNHOALFTHHYROYKNPDETPREGKTLTPCCRIKAKGEF TSEFDSMVAAPARSLSLHRRPPLLRRAITDALOGKPFHIDPLNRRQSTNTLAI ECGLATSESGNLSITRAIRALKEFSELDITVOTDPOICNITFDITFTALSA LVDSDVAVAARRSRVEMENOQREKQRLDEDLIARAFREFRFRFSYOTERRAH GLKRAARADVDVTRRDIDAIIVRQLFREIAEGRFVGNDAVREKARKYKEMLMASR NNNYTRLATGAT"
 1741. .2935
 /note="oriR"
 complement(3427, .3645)
 /gene="Y0006"
 complement(3427, .3645)
 /gene="Y0006"
 /note="OriS (f72): 42 pct identical (0 gaps) to 33 residues of an approx. 216 aa protein GENEPEPT: g1|2055297, Imp2 Xenopus laevis"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAC69763.1"
 /db_xref="GI:3822043"
 /translation="MRSPLAYGCSYTVVWVYQLHNDIHKSAVKHRLYHVLGQASS AQPVEFMPHRCMKAPADIVNSYDNRR"
 4758. .5186
 /gene="Y0008"
 4758. .5186
 /gene="Y0008"
 /note="OriT (o142): 31 pct identical (1 gap) to 48 residues of an approx. 104 aa protein GENEPEPT: g1|2149940, ori1 P. syringae"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAC69764.1"
 /db_xref="GI:3822044"
 /translation="MINTFETELPKIASFGDLKLSODEYGCCELLINDRYVIMRA DELINLTLLGPIIGSGEARSASQLEFCYSINLKNDCGCFANSEELGLIAPNHL SIDLNVENVSKRIANFYDMLSVSLPAETAPSYISTIG"
 5204. .7402
 /gene="ypkA"
 /note="Y0009"
 5204. .7402
 /gene="ypkA"

	Db	38099	AATGAGTAATTCTGATGTATATCGAATTCGACGAAGAATACTTAGCTTATTTTCTAACCC	38156
Oy	826	GAGATACCATTCTTTAAAGCGGCTCATTTATGCACAACCACTGCCAAAATGGCATCMAAGCA	885	
Db	38159	GAGATGCCATTCTTTAAAGCGGCTCATTTATGCACAACCACTGCCAAAATGGCATCMAAGCA	38218	
Oy	886	GTAANAAGTCTCTTAATCATATCGSCGAATATCACAAATGGSAATGGGGGCTTCATGGCA	945	
Db	38219	GTAANAAGTCTCTTAATCATATCGSCGAATATCACAAATGGSAATGGGGGCTTCATGGCA	38278	
Oy	946	GTAATGATTTCTCTTTAAACCGCGCATGTATCATGTATGATTTATTTGAAAGATTTGTT	1005	
Db	38279	GTAATGATTTCTCTTTAAACCGCGCATGTATCATGTATGATTTATTTGAAAGATTTGTT	38338	
Oy	1006	GATTCATGTAATCATATGATGGTGATGCCGCTAGACAAATTGGCTGAAGAATTAGCTGAGCTT	1065	
Db	38339	GATTCATGTAATCATATGATGGTGATGCCGCTAGACAAATTGGCTGAAGAATTAGCTGAGCTT	38399	
Oy	1066	ACCGCGCAATTAAGATTTATTCAGATTAATTCAGATTCAGACCGCAAAATTAATAGCATCTGTACT	1125	
Db	38399	ACCGCGCAATTAAGATTTATTCAGATTAATTCAGATTCAGACCGCAAAATTAATAGCATCTGTACT	38458	
Oy	1126	AGTGGCACCAATATATTCATGATTAATTCATTAATTCATGATGATTAATAATTAATAGG	1185	
Db	38459	AGTGGCACCAATATATTCATGATTAATTCATTAATTCATGATGATTAATAATTAATAGG	38518	
Oy	1186	TATACAGATGAAGAGATTTTTTAAAGCCAGCGCAGATCAAAATTTCTCGAAAAATGCGT	1245	
Db	38519	TATACAGATGAAGAGATTTTTTAAAGCCAGCGCAGATCAAAATTTCTCGAAAAATGCGT	38578	
Oy	1246	CAATCACCACTTAGGTGGATGGAGCGAGCAAAAAAATAGTCTCGATAAAGACTTTCTT	1305	
Db	38579	CAATCACCACTTAGGTGGATGGAGCGAGCAAAAAAATAGTCTCGATAAAGACTTTCTT	38638	
Oy	1306	GGAAGTGAAGATTAAGAAGACCGGGCGTTGGGTATCTGAAAAACCTCATCTCTTAATAT	1365	
Db	38639	GGAAGTGAAGATTAAGAAGACCGGGCGTTGGGTATCTGAAAAACCTCATCTCTTAATAT	38699	
Oy	1366	AAAGATTAATGAATTAATCTCACTTTGGCACACACTGCTCGATGAATGATCAAGCGCGTC	1425	
Db	38699	AAAGATTAATGAATTAATCTCACTTTGGCACACACTGCTCGATGAATGATCAAGCGCGTC	38758	
Oy	1426	AACGACTTGGTTAGCCAAAAAACAACTCAGCTGTCTGATATTACATCAGTTTTAATTA	1485	
Db	38759	AACGACTTGGTTAGCCAAAAAACAACTCAGCTGTCTGATATTACATCAGTTTTAATTA	38818	
Oy	1486	GCTATTAAACACTGAACGGTTCAATTCAGAAATATGATTAAGATGCAACGCTCGCA	1545	
Db	38819	GCTATTAAACACTGAACGGTTCAATTCAGAAATATGATTAAGATGCAACGCTCGCA	38878	
Oy	1546	GATGACACGCTCTGTGAATGA	1566	
Db	38879	GATGACACGCTCTGTGAATGA	38899	
RESULT	7			
LOCUS	A46411	1014 bp	DNA	PAT
DEFINITION	Sequence 1 from Patent W0524475.			07-MAR-1997
VERSION	A46411			
KEYWORDS	A46411.1 GI:2300612			
SOURCE	Yersinia pestis.			
ORGANISM	Bacteris; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
REFERENCE	1 (bases 1 to 1014)			
AUTHORS	Filball,R.W., Williamson,E.D. and Leary,S.E.			
TITLE	VACCINE COMPOSITIONS			
JOURNAL	Patent: WO 9524475-A 1 14-SEP-1995;			
COMMENT	SECUR DEFENCE BRIT (GB)			
FEATURES	Other publication AU 1653995 950925.			
	Location/Qualifiers			

source
1. 1014
/organism="Yersinia pestis"
/db_xref="taxon:632"
1. 990
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAA02893.1"
/db_xref="GI:2300613"

BASE COUNT 346 a 181 c 201 g 286 t
ORIGIN
KNIDISIKYDPRKSEVFAFANVITDDILELKKILATYLPEDAILKGHTYDQJONGIK
RVKEFLSSPTQWELRAFMVMSLTADRIDDLKIVDSNNHGDASKIREL
AELTAEIKIVSIOAEIKNHLSSGGINIHDKSINLMDKNLYGYTDEIFKASAEYKI
LEKMPOTIIVDGESEKIVSIKDFLSEKRNKRGALGNKNSYXKNDNELSHFATTC
SDKSRPINDLVSOXKTTLOSDITSRPNSEIENLNFIQKYSVMGRLLDDTSGR"

Query Match 62.4%; Score 976.8; DB 5; Length 1014;
Best Local Similarity 99.8%; Pred. No. 1.8e-201;
Matches 978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 587 TGATTAGAGCCTAGACAAAACCAACACATTTATTGAGATCTAGAAAAAGTTAGCG 646
11 TCATTAGAGCCTAGACAAAACCAACACATTTATTGAGATCTAGAAAAAGTTAGCG 70
647 TGGACCACTTACTGTCATGTTCTTCTGATTAGAGAATTGGTTCAGTTAGTCAAG 706
71 TGGACCACTTACTGTCATGTTCTTCTGATTAGAGAATTGGTTCAGTTAGTCAAG 130
707 ATAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTCGGAGTTTGCA 766
131 ATAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTCGGAGTTTGCA 190
767 ATAGATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 826
191 ATAGATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 250
827 AGGATACCATTTCTTAAAGCGGTCATATGACACCAACTGCAGAAATGCGATCAAGCGAG 886
251 AGGATACCATTTCTTAAAGCGGTCATATGACACCAACTGCAGAAATGCGATCAAGCGAG 310
887 TAAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 946
311 TAAAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 370
947 TAATCATTTCTTCTTAAAGCGGTCATGATGATGATGATGATGATGATGATGATG 1006
371 TAATCATTTCTTCTTAAAGCGGTCATGATGATGATGATGATGATGATGATGATG 430
1007 ATTCATGATATCATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
431 ATTCATGATATCATGATGATGATGATGATGATGATGATGATGATGATGATG 490
1067 CGCGCAATTAAGATTTATAGTATTCAGCCGAATTAAGATTTATAGATCTGCTAGTA 1126
491 CGCGCAATTAAGATTTATAGTATTCAGCCGAATTAAGATTTATAGATCTGCTAGTA 550
1127 GTGGACCAATAATATCCATGATTAATCTCATGATTAATAAATTTATATAGTT 1186
551 GTGGACCAATAATATCCATGATTAATCTCATGATTAATAAATTTATATAGTT 610
1187 ATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1246
611 ATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 670
1247 AAACCACTTACTGTCATGTTCTTCTGATTAGAGAATTGGTTCAGTTAGTCAAG 1306
671 AAACCACTTACTGTCATGTTCTTCTGATTAGAGAATTGGTTCAGTTAGTCAAG 130
1307 GAATGAGATTAAGAAACCGGGGCTTGGTAACTGAAAACTCATCTTTATATA 1366
|||||

Db 731 GAATGAGATTAAGAAACCGGGGCTTGGTAACTGAAAACTCATCTTTATATA 790
QY 1367 AAGATAATTAAGATTAATCTCATGTTGCGACCACTGTCGATTAAGCCAGCGCTCA 1426
Db 791 AAGATAATTAAGATTAATCTCATGTTGCGACCACTGTCGATTAAGCCAGCGCTCA 850
QY 1427 ACGACTTGGTTAGCCAAAACCACTGCTGTCATGATTAATCATCAGTTTAAATTCAG 1486
Db 851 ACGACTTGGTTAGCCAAAACCACTGCTGTCATGATTAATCATCAGTTTAAATTCAG 910
QY 1487 CTATTGAAGCACTGACCGTTTCATGAGAAATATGATGATGATGATGATGATG 1546
Db 911 CTATTGAAGCACTGACCGTTTCATGAGAAATATGATGATGATGATGATGATG 970
QY 1547 ATGACACGCTGCTGTAATGA 1566
Db 971 ATGACACGCTGCTGTAATGA 990

RESULT 8
A56793 1014 bp DNA PAT 03-MAR-1998
LOCUS A56793 Sequence 1 from Patent WO9628551.
DEFINITION A56793
ACCESSION A56793
VERSION A56793.1 GI:3712808
KEYWORDS
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Tiddell, R.W., Williamson, E.D., Leary, Sophie, E.C., Oyston, P.C.,
Bennett and Alice, M.
TITLE VACCINES FOR PLAGUE
JOURNAL Patent: NO 9628551-A 1 19-SEP-1996;
COMMENT SECR DEFENCE (GB)
other publication ZA 9602036 960716
other publication AU 4951196 961002.
FEATURES
location/Qualifiers
source 1. 1014
/organism="Yersinia pestis"
/db_xref="taxon:632"
1. 990
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAA03419.1"
/db_xref="GI:4530021"

BASE COUNT 346 a 181 c 201 g 286 t
ORIGIN
KNIDISIKYDPRKSEVFAFANVITDDILELKKILATYLPEDAILKGHTYDQJONGIK
RVKEFLSSPTQWELRAFMVMSLTADRIDDLKIVDSNNHGDASKIREL
AELTAEIKIVSIOAEIKNHLSSGGINIHDKSINLMDKNLYGYTDEIFKASAEYKI
LEKMPOTIIVDGESEKIVSIKDFLSEKRNKRGALGNKNSYXKNDNELSHFATTC
SDKSRPINDLVSOXKTTLOSDITSRPNSEIENLNFIQKYSVMGRLLDDTSGR"

Query Match 62.4%; Score 976.8; DB 5; Length 1014;
Best Local Similarity 99.8%; Pred. No. 1.8e-201;
Matches 978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 587 TGATTAGAGCCTAGACAAAACCAACACATTTATTGAGATCTAGAAAAAGTTAGCG 646
11 TCATTAGAGCCTAGACAAAACCAACACATTTATTGAGATCTAGAAAAAGTTAGCG 70
647 TGGACCACTTACTGTCATGTTCTTCTGATTAGAGAATTGGTTCAGTTAGTCAAG 706
71 TGGACCACTTACTGTCATGTTCTTCTGATTAGAGAATTGGTTCAGTTAGTCAAG 130
707 ATAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTCGGAGTTTGCA 766
131 ATAAATATAGATATTTCCATTAAATATGATCCAGAAAAGATTCGGAGTTTGCA 190

OY	767	ATAGAGTATTACTGATGTATGCATATGGCTCAAGAAAACCTTAGCTATTATTTCTACCCG	826
Dd	191	ATAGCATATTACTGATGTATGCATATGGCTCAAGAAAACCTTAGCTATTATTTCTACCCG	250
OY	827	AGGATACCATTCCTTAAGAAGCGGTCATTTATGACAACCACACTGCAAAATGGCATCAAGCGAG	886
Dd	251	AGGATGCCATTCCTTAAGAAGCGGTCATTTATGACAACCACACTGCAAAATGGCATCAAGCGAG	310
OY	887	TAAAGAGATTCCTTGATATATATGCGCGAATACACAAATGGGAATTTGCGGGCTTCATGGCAG	946
Dd	311	TAAAGAGATTCCTTGATATATATGCGCGAATACACAAATGGGAATTTGCGGGCTTCATGGCAG	370
OY	947	TATGCAATTCCTCTTAACCGCGGATCGATATGATGATATTTTGAAGATGTTG	1006
Dd	371	TATGCAATTCCTCTTTAACCGCGGATCGATATGATGATATTTTGAAGATGTTG	430
OY	1007	ATTCAATGATCATCATGCTGATGCGCGTAGCACAGTTGCGTGGAAGATTAGCTGACCTTA	1066
Dd	431	ATTCAATGATCATCATGCTGATGCGCGTAGCACAGTTGCGTGGAAGATTAGCTGACCTTA	490
OY	1067	CGCGCAATTAAGAATTTATTCAGTATTCACACCGCAATATATAGCATCTCTTAGTA	1126
Dd	491	CGCGCAATTAAGAATTTATTCAGTATTCACACCGCAATATATAGCATCTCTCTTAGTA	550
OY	1127	GTCGCACCATTAATATCCATGATATAATCCATTAATCTCATGATGATAAAAAATTATATGGT	1186
Dd	551	GTCGCACCATTAATATATCCATGATATAATCCATTAATCTCATGATGATAAAAAATTATATGGT	610
OY	1187	ATACAGATGATAGAGATTTTTTAAAGCCAGCGCAGAGTAGTACAAAATTCTCGAGAAAAATGCTC	1246
Dd	611	ATACAGATGATAGAGATTTTTTAAAGCCAGCGCAGAGTAGTACAAAATTCTCGAGAAAAATGCTC	670
OY	1247	AATACCAACCATTCAGGTGATGGGAGGAGGAAAAAATAGCTGATTAAGGAACTTCTTG	1306
Dd	671	AATACCAACCATTCAGGTGATGGGAGGAGGAAAAAATAGCTGATTAAGGAACTTCTTG	730
OY	1307	GAAGTGTAGATTAAGAACACCGGGCGTGGGTATCTGAAAAACTATACTCTTATATA	1366
Dd	731	GAAGTGTAGATTAAGAACACCGGGCGTGGGTATCTGAAAAACTATACTCTTATATA	790
OY	1367	AAGATTAATATGATTAATCTCACTTTGCCACACACTGCTCGGATGAAGTCCAGGCGCCTCA	1426
Dd	791	AAGATTAATATGATTAATCTCACTTTGCCACACACTGCTCGGATGAAGTCCAGGCGCCTCA	850
OY	1427	ACGACTGTGTAGCCAAAAACAACACTCAGGTGCTGATATTAATACATACGTTTAATTCAG	1486
Dd	851	ACGACTGTGTAGCCAAAAACAACACTCAGGTGCTGATATTAATACATACGTTTAATTCAG	910
OY	1487	CATATGAACACTGAAACCGTTTCATTCAGAAATATGATTCAGTATGATCAACGCTCTGAG	1546
Dd	911	CATATGAACACTGAAACCGTTTCATTCAGAAATATGATTCAGTATGATCAACGCTCTGAG	970
OY	1547	ATGACACGCTCTGTAAATGA 1566	
Dd	971	ATGACACGCTCTGTAAATGA 990	
RESULT	9		
LOCUS	A46413	1014 bp DNA	PAT 07-MAR-1997
DEFINITION	Sequence 3 from Patent WO9524475.		
ACCESSION	A46413		
VERSION	A46413.1	GI:2300614	
KEYWORDS			
SOURCE	Yersinia pestis.		
ORGANISM	Yersinia pestis; Proteobacteria; gamma subdivision; Enterobacteriaceae; Yersinia.		
REFERENCE	1 (bases 1 to 1014)		
AUTHORS	Titball, R.W., Williamson, E.D. and Leary, S.E.		
TITLE	VACCINE COMPOSITIONS		
JOURNAL	Patent: WO 9524475-A 3 14-SEP-1995;		

COMMENT	SECURITY DEFENCE BRIT (GB)
Other publication AU 1853995 950925.	
FEATURES	Location/Qualifiers
Source	1..1014
	/organism="Yersinia pestis"
	/db_xref="taxon:632"
	1..987
	/note="unamed protein product; protein sequence is in conflict with the conceptual translation"
	/codon_start=1
	/transl_table=1
	/protein_id="CAA02894.1"
	/db_xref="GI:2300615"
	/translation="GIRGIRAYEONPOHFEDELEKVPEDLTGHSVLEIYOLKDKNDISIKYDPRKDSYFANRVITDDLELKLITATFLPEDALIKGSHDNOLONGIKRVKEFLSSPSYDOWELNFKHVAVMHFSITLAIORIDDDILKTVDSNMNHGDARSLRELALTELKIKSVTIOAEINLHSSGITLNDKSLNMDKLVGYTEBEIFRASAERKILEKRPOTIYDGSSEKIVSIKDELISENKRGTGLNKLKSYSYNDNNLSFPAATCSKSRPINDLVSOKTTLOSDITSFNSNIAELNRFIOKIVSWORLLDDISGR"
BASE COUNT	343 a 184 c 205 g 282 t
ORIGIN	
Query Match	62.2%; Score 974.8; DB 5; Length 1014;
Best Local Similarity	99.8%; Pred. No. 4.8e-201;
Matches 976; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
589 ATTAGAGCTACGACAAACCAACACATTTTATTGAGATCTAGAAAAAGTTAGGGTG	648
13 ATTGAGGCTACGACAAACCAACACATTTTATTGAGATCTAGAAAAAGTTAGGGTG	72
649 GAACAACTTACTGGTCATAGTCTTCACTTTTAGAAGATTGGTCACTAGTCAAGAT	708
73 GAACAACTTACTGGTCATAGTCTTCACTTTTAGAAGATTGGTCACTAGTCAAGAT	132
709 AAAATATAGATTTCCATTAATATATATGCCAAGAAAGTTGGAGGTTTGGCAAT	768
133 AAAATATAGATTTCCATTAATATATATGCCAAGAAAGTTGGAGGTTTGGCAAT	192
769 AGAGTAAATTTACTGATGATGCAATTCGTCAAGAAATTCCTAGCTTATTTCTACCGGAG	828
139 AGAGTAAATTTACTGATGATGCAATTCGTCAAGAAATTCCTAGCTTATTTCTACCGGAG	252
829 GATACCAATCTTAAAGCGGGTCATTATGACAAACCAACTGCAAAATGGCATCAACGAGTA	888
253 GATGCCATTTCTTAAAGCGGGTCATTATGACAAACCAACTGCAAAATGGCATCAACGAGTA	312
889 AAAGATTCCTTGATTCATGCCGAAATACAAATGGGAATTCGGGGGCTTCATGGCAGTA	948
313 AAAGATTCCTTGATTCATGCCGAAATACAAATGGGAATTCGGGGGCTTCATGGCAGTA	372
949 ATGCATTTCTCTTTAACCGCGGATGATGCAATGATGATTTTGAAGAATGATTGGTAT	1008
373 ATGCATTTCTCTTTAACCGCGGATGATGCAATGATGATTTTGAAGAATGATTGGTAT	432
1009 TCAATGAATCATCATGATGATGCCGTAGCAAGTTGCGTGAAGAATTAGCTGAGCTTACC	1068
433 TCAATGAATCATCATGATGATGCCGTAGCAAGTTGCGTGAAGAATTAGCTGAGCTTACC	492
1069 GCCGAATTAAGATTTATTCAGTATTTTAAAGCGGAATTAATAAGCATCTCTAGTAGT	1128
493 GCCGAATTAAGATTTATTCAGTATTTTAAAGCGGAATTAATAAGCATCTCTAGTAGT	552
1129 GGCACCAATTAATTCATGATGATTAATCCATTATGTCATGATGATTAATAATTTATGGTTAT	1188
553 GGCACCAATTAATTCATGATGATTAATCCATTATGTCATGATGATTAATAATTTATGGTTAT	612
1189 ACAGATGAAGAGATTTTAAAGCCAGCCCAAGTACAAAATTCCTGAGAAAATGCTCTCA	1248
613 ACAGATGAAGAGATTTTAAAGCCAGCCCAAGTACAAAATTCCTGAGAAAATGCTCTCA	672
1249 ACCACCAATTCAGGTGATGGAGCGAGAAAAAATAGTCTGATTAAGGACTTTCTTGA	1308

Db 673 ACCACATTCAGTGGATGGAGCGAGAAAAATAGTCTCGATAAAGACCTTCTTGA 732
Oy 1309 AGTGAGATTAAGAACCGGGCGTTGGTAACTGAAAACTACTCTATATATAA 1368
Db 733 AGTGAGATTAAGAACCGGGCGTTGGTAACTGAAAACTACTCTATATATAA 792
Oy 1369 GATAAATAATGATATATCTACTTTCGCAACACCTGCTGGATAGTCAGGCCGCTCAAC 1428
Db 793 GATAAATAATGATATATCTACTTTCGCAACACCTGCTGGATAGTCAGGCCGCTCAAC 852
Oy 1429 GACTTGGTTAGCAAAAAACAACACTAGCTGTCTGATATTAATCAACGTTTAACTCAGCT 1488
Db 853 GACTTGGTTAGCAAAAAACAACACTAGCTGTCTGATATTAATCAACGTTTAACTCAGCT 912
Oy 1489 ATTGAAGACTGAACCGTTTCAATGAAATGATTCAGATGATGCAACGCTGCTAGT 1548
Db 913 ATTGAAGACTGAACCGTTTCAATGAAATGATTCAGATGATGCAACGCTGCTAGT 972
Oy 1549 GACACGCTGTGTAATGA 1566
Db 973 GACACGCTGTGTAATGA 990

RESULT 10
LOCUS A56795 1014 bp DNA
DEFINITION Sequence 3 from Patent WO9628551.
ACCESSION A56795
VERSION A56795.1 GI:3712810
KEYWORDS
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Tlball,R.W., Williamson,E.D., Leary, Sophie,E.C., Oyston,P.C.,
Bennett and Alice,M.
TITLE VACCINES FOR PLAGUE
JOURNAL Patent: WO 9628551-A 3 19-SEP-1996;
COMMENT SECR DEFENCE (GB) 2A 9602036 960716
Other publication AU 4951196 961002.
FEATURES
source location/Qualifiers
1. 1014
/organism="Yersinia pestis"
/db_xref="taxon:632"
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAA03420.1"
/db_xref="GI:3712811"
/translation="GIPGIRAYEONPQHFIEDLEKRVLEQJLTHGSSVLELYLVKQ
KNIDISIKYPRKDSYEVANFVTDIEDLKLAYFLPEDAIKGHYDQNLQNGIK
RVKELSSPTOMELRAFMAVHPSLADIDDDILKVIDSMNHGDASRKLREEL
AELTAEIKIVSIOAEINKHLSGGTINIHKSINMDKNKXTGDEIRFASAEYKI
LEKRPQTTIOVDSSEKIVSIKDELGSNKRITGALGNKNSIYNKDNENSHRTTC
SDKSRPLNDLVSOXTTQDLSITSRFNSIEALNRFIOKIDYSVMORLDDITSGK"

BASE COUNT 343 a 184 c 205 g 282 t
ORIGIN

Query Match 62.2%; Score 974.8; DB 5; Length 1014;
Best Local Similarity 99.8%; Pred. No. 4.8e-201;
Matches 976; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 589 ATTAGAGCTACGACAAACCAACACATTATTAGAGATGAGAAATTAAGGGG 648
Db 13 ATGAGAGCTACGACAAACCAACACATTATTAGAGATGAGAAATTAAGGGG 72
Oy 649 GACAACTTACTGTCATGCTTCTTCACTTTAGAGATGCTTCACTTCAAGAT 708
Db 73 GACAACTTACTGTCATGCTTCTTCACTTTAGAGATGCTTCACTTCAAGAT 132

Oy 709 AAAATATAGATATTTCCATTAAATATATATCCAGAAAAAGATGGAGTTTGGCAAT 768
Db 133 AAAATATAGATATTTCCATTAAATATATATCCAGAAAAAGATGGAGTTTGGCAAT 192
Oy 769 AGAGTAATTAAGATGATATGAAATGCTCAAGAAATCTAGCTTATTTCTACCCGAG 828
Db 193 AGAGTAATTAAGATGATATGAAATGCTCAAGAAATCTAGCTTATTTCTACCCGAG 252
Oy 829 GATACATTTCTTAAAGCGGCTCATTTATGACAAACCACTGCAAAATGGCATCAAGCGATA 888
Db 253 GATGCAATTTCTTAAAGCGGCTCATTTATGACAAACCACTGCAAAATGGCATCAAGCGATA 312
Oy 889 AAAGATCTCTGATATATATCCCGAATACAAAGGGAATGGCGGCTTATGCGACATA 948
Db 313 AAAGATCTCTGATATATATCCCGAATACAAAGGGAATGGCGGCTTATGCGACATA 372
Oy 949 ATGATTTCTCTTTTAAACCGCGATCGTATCGATGATGATTTTGAAGTATGTTGAT 1008
Db 373 ATGATTTCTCTTTTAAACCGCGATCGTATCGATGATGATTTTGAAGTATGTTGAT 432
Oy 1009 TCAATGAATCATCATGATGATGATGCCGTAGCAAGTTGCGTGAAGAAATGCTGAGCTTACC 1068
Db 433 TCAATGAATCATCATGATGATGATGCCGTAGCAAGTTGCGTGAAGAAATGCTGAGCTTACC 492
Oy 1069 GCCGAATTAAGATTTATTTCACTGATTTTCAAGCCGAAATTAATAGCATCTGCTAGTACT 1128
Db 493 GCCGAATTAAGATTTATTTCACTGATTTTCAAGCCGAAATTAATAGCATCTGCTAGTACT 552
Oy 1129 GGCACATTAATATCATGATATTAATCATTAATCATGATGATTAATTAATGATGTTAT 1188
Db 553 GGCACATTAATATCATGATATTAATCATTAATCATGATGATTAATTAATGATGTTAT 612
Oy 1189 ACAGATGAAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTGAGAAAAATGCTCAA 1248
Db 613 ACAGATGAAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTGAGAAAAATGCTCAA 672
Oy 1249 ACCACATTCAGTGGATGGAGCGAGGAAAAAATAGCTGATTAAGACTTTCTTGA 1308
Db 673 ACCACATTCAGTGGATGGAGCGAGGAAAAAATAGCTGATTAAGACTTTCTTGA 732
Oy 1309 AGTGAGATTAAGAACCGGGCGTTGGTAACTGAAAACTACTCTATATATAA 1368
Db 733 AGTGAGATTAAGAACCGGGCGTTGGTAACTGAAAACTACTCTATATATAA 792
Oy 1369 GATAAATAATGATATATCTACTTTCGCAACACCTGCTGGATAGTCAGGCCGCTCAAC 1428
Db 793 GATAAATAATGATATATCTACTTTCGCAACACCTGCTGGATAGTCAGGCCGCTCAAC 852
Oy 1429 GACTTGGTTAGCAAAAAACAACACTAGCTGTCTGATATTAATCAACGTTTAACTCAGCT 1488
Db 853 GACTTGGTTAGCAAAAAACAACACTAGCTGTCTGATATTAATCAACGTTTAACTCAGCT 912
Oy 1489 ATTGAAGACTGAACCGTTTCAATGAAATGATTCAGATGATGCAACGCTGCTAGT 1548
Db 913 ATTGAAGACTGAACCGTTTCAATGAAATGATTCAGATGATGCAACGCTGCTAGT 972
Oy 1549 GACACGCTGTGTAATGA 1566
Db 973 GACACGCTGTGTAATGA 990

RESULT 11
LOCUS AF167310 981 bp DNA
DEFINITION Yersinia pestis strain Angola V antigen (IcrV) gene, complete cds.
ACCESSION AF167310
VERSION AF167310.1 GI:7578514
KEYWORDS
SOURCE Yersinia pestis.
ORGANISM Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.

REFERENCE 1 (bases 1 to 981)
 AUTHORS Adair, D.M., Worsham, P.L., Hill, K.K., Klevytska, A.M., Jackson, P.J.,
 TITLE Friedlander, A.M. and Keim, P.
 JOURNAL J. Clin. Microbiol. 38 (4), 1516-1519 (2000)
 MEDLINE 20211685
 REFERENCE 2 (bases 1 to 981)
 AUTHORS Hill, K.K. and Jackson, P.J.
 TITLE Direct Submission
 JOURNAL Submitted (08-JUL-1999) Bioscience Division, Los Alamos National
 Laboratory, MS:M888, Los Alamos, NM 87545, USA

FEATURES
 source
 1. 961
 /organism="Yersinia pestis"
 /strain="angola"
 /db_xref="taxon:632"
 1. 981
 /gene="lcrV"
 1. 981
 /gene="lcrV"
 /codon_start=-1
 /transl_table=11
 /product="Y antigen"
 /protein_id="AA064077.1"
 /db_xref="GI:7578513"
 /translation="MIRAYEQNPQHFIEDLENYRVQLTGHSSVLELYQVKKRNI
 DISIKYDPKDESEVFNARYITDDIELRKLAFEDLALKGNHONOLNIRK
 EFLESPNTOELRFAFVHFSLEDRIDDLIKYIDSMNHGARSRLRELEL
 TAEIKIYVIOAEINKHLSSTSGTINHDKSNIMDNKYVYDDELFKSAEYKILEK
 MPTQITQVDSSEKIVSIRKDFLSESKRTRALDNKSYKNDNNELSPFATTSQK
 SRPLNDVDSQKTTLDSPITSFNSAIEALNRFIQKDYDMRLLDPTSGK"

BASE COUNT 334 a 175 c 196 g 276 t
 ORIGIN

Query Match 62.2%; Score 974.6; DB 2; Length 981;
 Best Local Similarity 99.6%; Pred. No. 5.4e-201;
 Matches 977; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 586 ATGATTAGAGCCTAGCAACAAACCCACACATTTTATGAGATCTAGAAAAGTAGG 645
 |||||||
 DB 1 ATGATTAGAGCCTAGCAACAAACCCACACATTTTATGAGATCTAGAAAAGTAGG 60
 |||||||

QY 646 GTGGACAACCTACTGCTGATGTTCTTCACTTTAGAGAATGGTTCACTAGTCAAA 705
 |||||||
 DB 61 GTGGACAACCTACTGCTGATGTTCTTCACTTTAGAGAATGGTTCACTAGTCAAA 120
 |||||||

QY 706 GATAAAATATAGATATTTCCATTAATATGATCCCAAGAAAAGTTGGAGGTTTGGC 765
 |||||||
 DB 121 GATAAAATATAGATATTTCCATTAATATGATCCCAAGAAAAGTTGGAGGTTTGGC 180
 |||||||

QY 766 AATGAGTAATTAATGATGATGATGATGATGCTCAAGAAAATCTAGCTATTTCTACCC 825
 |||||||
 DB 181 AATGAGTAATTAATGATGATGATGATGATGCTCAAGAAAATCTAGCTATTTCTACCC 240
 |||||||

QY 826 GAGGATACCATTTCTTAAAGGGGTCATTATGACAAACCACTGCAAAATGGCATCAACGA 885
 |||||||
 DB 241 GAGGATACCATTTCTTAAAGGGGTCATTATGACAAACCACTGCAAAATGGCATCAACGA 300
 |||||||

QY 886 GTAAAAAGTCTCTTGAATCATCTGCGGATACAAATGGAATGGCGGCTTATGGCA 945
 |||||||
 DB 301 GTAAAAAGTCTCTTGAATCATCTGCGGATACAAATGGAATGGCGGCTTATGGCA 360
 |||||||

QY 946 GTATGATTTCTCTTAAACCGCGATGCTGATGATGATGATTTTGAAGTATGTT 1005
 |||||||
 DB 361 GTATGATTTCTCTTAAACCGCGATGCTGATGATGATGATTTTGAAGTATGTT 420
 |||||||

QY 1006 GATTCAATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
 |||||||
 DB 421 GATTCAATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 |||||||

QY 1066 ACCGCCAATTAAGATTATTCAGTTATTCAGCCGAATTAATGAATGATCTGTACT 1125
 |||||||

DB 481 ACCGCCAATTAAGATTATTCAGTTATTCAGCCGAATTAATGAATGATCTGTACT 540
 |||||||

QY 1126 AGTGGACATTAATATTCATGATTAATTCATTAATTCATGATTAATTAATTAATGCT 1185
 |||||||

DB 541 AGTGGACATTAATATTCATGATTAATTCATTAATTCATGATTAATTAATTAATGCT 600
 |||||||

QY 1186 TATACAGATGATGATGATTTTAAAGCCAGGACAGATCAAAATTCGAGAAAATGCT 1245
 |||||||

DB 601 TATACAGATGATGATGATTTTAAAGCCAGGACAGATCAAAATTCGAGAAAATGCT 660
 |||||||

QY 1246 CAATCCACATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
 |||||||

DB 661 CAATCCACATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 |||||||

QY 1306 GGAAGTGAATTAAGAACCGGCGCTTGGTATCTGAAAACATCATCTTATAT 1365
 |||||||

DB 721 GGAAGTGAATTAAGAACCGGCGCTTGGTATCTGAAAACATCATCTTATAT 780
 |||||||

QY 1366 AAGATTAATTAAGATTAATTCATCTTGGCACACCTGCTGGATAAGTCCAGCGCTC 1425
 |||||||

DB 781 AAGATTAATTAAGATTAATTCATCTTGGCACACCTGCTGGATAAGTCCAGCGCTC 840
 |||||||

QY 1426 AAGACTTGGTACCCAAAACACACTGCTGCTGATATTCATCAGCTTTATTC 1485
 |||||||

DB 841 AAGACTTGGTACCCAAAACACACTGCTGCTGATATTCATCAGCTTTATTC 900
 |||||||

QY 1486 GCTATTGACACTGAAACCTTTTCATGAGAAATATGATGATGATGATGATGATGAT 1545
 |||||||

DB 901 GCTATTGACACTGAAACCTTTTCATGAGAAATATGATGATGATGATGATGATGAT 960
 |||||||

QY 1546 GATGACACGTCGTGTAATGA 1566
 |||||||

DB 961 GATGACACGTCGTGTAATGA 981
 |||||||

RESULT 12
 AF167309 975 bp DNA BCT 17-APR-2000
 LOCUS Yersinia pestis strain Pestoides F V antigen (lcrV) gene, complete
 DEFINITION
 ACCESSION AF167309
 VERSION AF167309.1 GI:7578512
 KEYWORDS
 SOURCE Yersinia pestis.
 ORGANISM Yersinia pestis.
 Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.

REFERENCE 1 (bases 1 to 975)
 AUTHORS Adair, D.M., Worsham, P.L., Hill, K.K., Klevytska, A.M., Jackson, P.J.,
 TITLE Friedlander, A.M. and Keim, P.
 JOURNAL J. Clin. Microbiol. 38 (4), 1516-1519 (2000)
 MEDLINE 20211685
 REFERENCE 2 (bases 1 to 975)
 AUTHORS Hill, K.K. and Jackson, P.J.
 TITLE Direct Submission
 JOURNAL Submitted (08-JUL-1999) Bioscience Division, Los Alamos National
 Laboratory, MS:M888, Los Alamos, NM 87545, USA

FEATURES
 source
 1. 975
 /organism="Yersinia pestis"
 /strain="Pestoides F"
 /db_xref="taxon:632"
 1. 975
 /gene="lcrV"
 1. 975
 /gene="lcrV"
 /codon_start=-1
 /transl_table=11
 /product="Y antigen"
 /protein_id="AA064076.1"
 /db_xref="GI:7578513"
 /translation="MIRAYEQNPQHFIEDLENYRVQLTGHSSVLELYQVKKRNI"

DISIKYDPRKSEVFANRYITDDIELKILAYEDELALKGHYDNOJONGIKREK
 EFLESPNTOEWELRAFMVAHFSLADRIDDDILKTYVSMNHGARSRLRELAEL
 TAEKITYIOAEINKHLSGGTINIHDKSINIKYTGTEEFKSAEYKILIEK
 MPOTIIOVGESEKRIYISINDFISEKRRGALGNLNSYSYNDNNELSHFATCSDK
 SRPLNDVOKTQOLSDITSRENSAEALNRFIOKIDSVORLDDTGR"

BASE COUNT 334 a 174 c 194 g 273 t
 ORIGIN

Query Match 61.8%; Score 967.4; DB 2; Length 975;
 Best Local Similarity 99.9%; Pred. No. 1.9e-199;

Matches 968; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

586 ATGATTAGACCTACGACAAACCCACACATTTATTGAGATCTAGAAAAAGTTAGG 645
 1 ATGATTAGACCTACGACAAACCCACACATTTATTGAGATCTAGAAAAAGTTAGG 60
 646 GTGGACAACTTACTGTCATGTTCTTCAGTTTAAAGAAATGGTTCAAGTCAAA 705
 61 GTGGACAACTTACTGTCATGTTCTTCAGTTTAAAGAAATGGTTCAAGTCAAA 120
 706 GATPAAAAATATAGTATTTCCATTAATATGATCCCAAGAAAGTTGGAGTTTGGC 765
 121 GATPAAAAATATAGTATTTCCATTAATATGATCCCAAGAAAGTTGGAGTTTGGC 180
 766 AATAGATAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 825
 181 AATAGATAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 826 GAGATACCACTTCTTAAAGCGGTCATTAAGACCAACCACTGCAAAATGGCATACGCA 885
 241 GAGATACCACTTCTTAAAGCGGTCATTAAGACCAACCACTGCAAAATGGCATACGCA 300
 886 GTAAAAAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 945
 301 GTAAAAAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 946 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1005
 361 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
 1006 GATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1065
 421 GATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 1066 ACCGCCCAATTAAGATTTATCTGTTATTCAGCCCAATTAATTAAGCATCTGTAGT 1125
 481 ACCGCCCAATTAAGATTTATCTGTTATTCAGCCCAATTAATTAAGCATCTGTAGT 540
 1126 AGTGGCACATTAATATCCATGATTAATCATGATGATGATGATGATGATGATGAT 1185
 541 AGTGGCACATTAATATCCATGATTAATCATGATGATGATGATGATGATGATGATG 600
 1186 TATACAGATGAAGATTTTAAAGCCAGCCAGATGACAAATTCGAGAAAAATGCTT 1245
 601 TATACAGATGAAGATTTTAAAGCCAGCCAGATGACAAATTCGAGAAAAATGCTT 660
 1246 CAACACCACTTCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1305
 661 CAACACCACTTCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
 1306 GGAAGTGAAGATTAAGAACCGGGGGGTGGTAACTGTAATGTAATGTAATGTAAT 1365
 721 GGAAGTGAAGATTAAGAACCGGGGGGTGGTAACTGTAATGTAATGTAATGTAAT 780
 1366 AAGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425
 781 AAGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 1426 AAGACCTGTTGAGCCAAAAACCACTGAGTGTCTGATATATACATCAGTTTAAATCA 1485
 841 AAGACCTGTTGAGCCAAAAACCACTGAGTGTCTGATATATACATCAGTTTAAATCA 900

QY 1486 GCATTGACCACTGACACCGTTTCATTCAGAAATATGATTCAGTATGACACGCTGCTA 1545
 DB 901 GCATTGACCACTGACACCGTTTCATTCAGAAATATGATTCAGTATGACACGCTGCTA 960
 QY 1546 GATGACACG 1554
 DB 961 GATGACACG 969

RESULT 13
 YPTPIVANT 981 bp DNA BCT 13-MAR-1997

LOCUS YPTPIVANT 981 bp DNA BCT 13-MAR-1997
 DEFINITION Y.pseudotuberculosis V antigen gene.
 ACCESSION X96802
 VERSION X96802.1 GI:1405834

KEYWORDS V antigen.
 SOURCE Yersinia pseudotuberculosis.
 ORGANISM Yersinia pseudotuberculosis
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Yersinia.

REFERENCE 1 (bases 1 to 981)
 Rogenkamp, A., Geiger, A.M., Leitritz, L., Kessler, A. and
 Heesemann, J.
 TITLE Passive immunity to infection with Yersinia spp. mediated by
 anti-recombinant V antigen is dependent on polymorphism of V
 antigen

JOURNAL Infect. Immun. 65 (2), 446-451 (1997)
 MEDLINE 97162308
 REFERENCE 2 (bases 1 to 981)
 Rogenkamp, A.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAR-1996) A. Rogenkamp, Hygiene and Microbiology,
 University of Wuerzburg, Josef-Schneider Strasse 2, 97080
 Wuerzburg, FRG

FEATURES
 source
 Location/Qualifiers
 1..981
 /organism="Yersinia pseudotuberculosis"
 /sub_species="type I"
 /db_xref="taxon:633"

CDS
 1..981
 /codon_start=1
 /transl_table=11
 /product="V antigen"
 /protein_id="CA65594.1"
 /db_xref="GI:1405835"
 /db_xref="SWISS-PROT:P23994"

ORIGIN
 1..981
 /translation="MIRAYEONPOHEIEDLEKRYEOLTGHSVLEELVQVYKDKNI
 DISIKYDPRKSEVFANRYITDDIELKILAYEDELALKGHYDNOJONGIKREK
 EFLESPNTOEWELRAFMVAHFSLADRIDDDILKTYVSMNHGARSRLRELAEL
 TAEKITYIOAEINKHLSGGTINIHDKSINIKYTGTEEFKSAEYKILIEK
 MPOTIIOVGESEKRIYISINDFISEKRRGALGNLNSYSYNDNNELSHFATCSDK
 SRPLNDVOKTQOLSDITSRENSAEALNRFIOKIDSVORLDDTGR"

BASE COUNT 341 a 174 c 191 g 275 t

Query Match 61.1%; Score 957; DB 1; Length 981;
 Best Local Similarity 98.5%; Pred. No. 3.5e-197;

Matches 966; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

586 ATGATTAGACCTACGACAAACCCACACATTTATTGAGATCTAGAAAAAGTTAGG 645
 1 ATGATTAGACCTACGACAAACCCACACATTTATTGAGATCTAGAAAAAGTTAGG 60
 646 GTGGACAACTTACTGTCATGTTCTTCAGTTTAAAGAAATGGTTCAAGTCAAA 705
 61 GTGGACAACTTACTGTCATGTTCTTCAGTTTAAAGAAATGGTTCAAGTCAAA 120
 706 GATPAAAAATATAGTATTTCCATTAATATGATCCCAAGAAAGATTCGAGGTTTGGC 765
 121 GATPAAAAATATAGTATTTCCATTAATATGATCCCAAGAAAGATTCGAGGTTTGGC 180
 766 AATAGATAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 825

```

|||||
Db 181 AATAGATATATCTGATATATCGATTGCTCAAGAAAATCTTACTTATTTTCAACC 240
OY 826 GAGATACCATCTTAAAGCGGCTCATTTATGACAACTGCAAAATGGCATCAAGCA 885
Db 241 GAGGATGCATCTTAAAGCGGCTCATTTATGACAACTGCAAAATGGCATCAAGCA 300
OY 886 GTAAAGAGCTCTTGAATCATGCGCAATACAAATGGGAATGGCGGCTCATGCA 945
Db 301 GTAAAGAGCTCTTGAATCATGCGCAATACAAATGGGAATGGCGGCTCATGCA 360
OY 946 GTAAATGATCTCTTAAAGCGGCTCATGATGATGATTTTGAATGATGCT 1005
Db 361 GTAAATGATCTCTTAAAGCGGCTCATGATGATGATTTTGAATGATGCT 420
OY 1006 GATTCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
Db 421 GATTCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 480
OY 1066 ACCGCGCAATTAAGATTTTATTCAGTTATTCAGCCGAATTAATAGCATCTGTCTAGT 1125
Db 481 ACCGCGCAATTAAGATTTTATTCAGTTATTCAGCCGAATTAATAGCATCTGTCTAGT 540
OY 1126 AGTGGACCATTAATATTCATGATTAATATTCATGATTAATATTCATGAT 1185
Db 541 GGTGGACCATTAATATTCATGATTAATATTCATGATTAATATTCATGAT 600
OY 1186 TATACAGATGAAGATTTTAAAGCGGCTGATGATGATGATGATGATGATGAT 1245
Db 601 TATACAGATGAAGATTTTAAAGCGGCTGATGATGATGATGATGATGATGAT 660
OY 1246 CAACCCACCATTCAGTGGATGGAGCGAGAAAAAATAGTCTCGATTAAGACCTTTCT 1305
Db 661 CAACCCACCATTCAGTGGATGGAGCGAGAAAAAATAGTCTCGATTAAGACCTTTCT 720
OY 1306 GGAATGAGAAATTAAGAAACCGGCGGCTGATTAATCTGATTAATCTGATTAAT 1365
Db 721 GGAATGAGAAATTAAGAAACCGGCGGCTGATTAATCTGATTAATCTGATTAAT 780
OY 1366 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1425
Db 781 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
OY 1426 AAGCATTTGTTAGCCAAACAACTGCTGCTGATTAATTAATTAATTAATTAAT 1485
Db 841 AAGCATTTGTTAGCCAAACAACTGCTGCTGATTAATTAATTAATTAATTAAT 900
OY 1486 GCTATTAAGCACTGACGCTTCAATGAAATATGATTAATTAATTAATTAAT 1545
Db 901 GCTATTAAGCACTGACGCTTCAATGAAATATGATTAATTAATTAATTAATTAAT 960
OY 1546 GATGACACCTCTGTAATGA 1566
Db 961 GATGACACCTCTGTAATGA 981

RESULT 14
YEPICRGVHP 2201 bp DNA BCT 26-APR-1993
LOCUS Yersinia pseudotuberculosis V-antigen (lcrG, lcrV, lcrH) genes,
DEFINITION Complete cds.
ACCESSION M57893
VERSION M57893.1 GI:155456
KEYWORDS V-antigen.
SOURCE Yersinia pseudotuberculosis (strain YPIII (PIBI) DNA.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
REFERENCE 1 (bases 1 to 2201)
AUTHORS Bergman,T., Hakansson,S., Forsberg,A., Norlander,L., Macellaro,A.,
Baechman,A., Boelgen,I. and Wolf-Watz,H.
TITLE Analysis of the V antigen lcrvH-yopD operon of Yersinia
pseudotuberculosis: Evidence for a regulatory role of lcrH and lcrV

```

```

JOURNAL J. Bacteriol. 173, 1607-1616 (1991)
MEDLINE 91154114
FEATURES
SOURCE
Location/Qualifiers
1..2201
/organism="Yersinia pseudotuberculosis"
/strain="YPIII (PIBI)"
/db_xref="taxon:633"
-35_signal
219..224
/gene="lcr operon"
219..258
/gene="lcr operon"
-10_signal
241..246
/gene="lcr operon"
252..258
/gene="lcr operon"
RBS
264..551
/gene="lcrG"
264..551
CDS
/gene="lcrG"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="V-antigen"
/protein_id="AA27644.1"
/db_xref="GI:153457"
/transl_table="MKSHFDEYDKTLKQAEALVDSHRAKLQENCADIGLTPAV
MKIPAGSAEIKPAERELDEIKRQERQPOHPHYDGKRRPKPTMGQII"
553..1533
/gene="lcrV"
553..1533
CDS
/gene="lcrV"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="V-antigen"
/protein_id="AA27645.1"
/db_xref="GI:153458"
/transl_table="MIRAYBONPOHFIEDLEKVEQLTGSSVLELYVOLYKDKNI
DISIKYPRDSEYFANRYITDDIELKILAYFLPDALIKGHYDNOGLGKRYK
EELSSYNTOEWELAFNAVHIFSLADRIDDLKIVDSNMNHGDAKSLRELAEL
TAEIKYVIOAEIKNLKSSGCTINIKSINMDKLVGTGDEIKAKAEKILEK
MPQTTIDBGETEKIVSIRKFLBEKRTGALGKDSISYNDNNLSHFATCSOK
SRPNDLVSOQTLDITSIRNSAIEALNRFQKXDSVMORLDTSGK"
1546..2052
/gene="lcrH"
1546..2052
CDS
/gene="lcrH"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="V-antigen"
/protein_id="AA27646.1"
/db_xref="GI:155459"
/transl_table="MQEETDQEQVLAESFLKGGTILAMLEISDTEQLYSLEF
NQYSGKEDAHKVFQALCVDDHYDSFFELGACQKAMQYDLAHSYSYGLIMIK
EPREPHAAECLOKGLAESEGLFLAOELIADKPEKELSTRVSSMLEAIRIKREM
EHECVDPNP"

BASE COUNT 714 a 450 c 485 g 552 t
ORIGIN

Query Match 60.9%; Score 953.8; DB 1; Length 2201;
Best Local Similarity 98.3%; Pred. No. 1.6e-196;
Matches 964; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
OY 586 ATGATTAGAGCTTACGACAAACCCACACATTTATGAGATCTAGAAAAGTAGG 645
Db 553 ATGATTAGAGCTTACGACAAACCCACACATTTATGAGATCTAGAAAAGTAGG 612
OY 646 GTGGAACACTTACGACGATGCTTCAATTTAGAGAAATGGCTACATTAGTAA 705
Db 613 GTGGAACACTTACGACGATGCTTCAATTTAGAGAAATGGCTACATTAGTAA 672

```


QY	706	GATAAATATGATATTTCCATTTAAATATGATCCAGAAAAGATTGCGAGGTTTTC	765
Db	673	GATAAATATGATATTTCCATTTAAATATGATCCAGAAAAGATTGCGAGGTTTTC	732
QY	766	AATAGAGTAATTTCTGATGATATCGAATTTGCAAGAAAATTCATGCTTATTTCTACCC	825
Db	723	AATAGAGTAATTTCTGATGATATCGAATTTGCAAGAAAATTCATGCTTATTTCTACCC	792
QY	826	GAGGATACCATTTCTTAAAGCGGCTATTATGCAACAACATGCAAAATGGCATCAAGCA	885
Db	793	GAGGATGCGCATTTCTTAAAGCGGCTATTATGCAACAACATGCAAAATGGCATCAAGCA	852
QY	886	GTAAGAAAGTTCCCTGGAATCATGCGCGGAATACACAATGGGAATTCGCGGCGTCAATGGCA	945
Db	853	GTAAGAAAGTTCCCTGGAATCATGCGCGGAATACACAATGGGAATTCGCGGCGTCAATGGCA	912
QY	946	GTAATGCAATTTCTCTTAAACCGCCGATGATGATGATATTTTGAATGATGATT	1005
Db	913	GTAATGCAATTTCTCTTAAACCGCCGATGATGATGATATTTTGAATGATGATT	972
QY	1006	GATTCAATGATCATCATGATGATGATCCCGGTACCAAGTTGCGTGAAGAAATTTAGCTGAGTT	1065
Db	973	GATTCAATGATCATCATGATGATGATCCCGGTACCAAGTTGCGTGAAGAAATTTAGCTGAGTT	1032
QY	1066	ACCGCGGAATTTAAATATTTATTCATTTATTCAGCCGCAAAATTAATAGCATGTGCTAGT	1125
Db	1033	ACCGCGGAATTTAAATATTTATTCATTTATTCAGCGTGAATTAATAGCATGTGCTAGT	1092
QY	1126	AGTGCCACCATTAATTCATGATTAATTCATTAATTCATGATGATTAATTAATGATGAT	1185
Db	1093	GGTGCGACCATTAATTCATGATTAATTCATTAATTCATGATGATTAATTAATGATGAT	1152
QY	1186	TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAAATTTCTGAGAAAATGCT	1245
Db	1153	TATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAAATTTCTGAGAAAATGCT	1212
QY	1246	CAAAACCACCATTCAGTGGATGGAGCGGAGGAGAAAATATGATCGATAAGACATTCTT	1305
Db	1213	CAAAACCACCATTCAGTGGATGGAGCGGAGGAGAAAATATGATCGATAAGACATTCTT	1272
QY	1306	GGAATGGAATTAAGAAACCGGGCGTTGGGTATTCGAAAATGCTCATCTTATAT	1365
Db	1273	GGAATGGAATTAAGAAACCGGGCGTTGGGTATTCGAAAATGCTCATCTTATAT	1332
QY	1366	AAAGATTAATGAATTAATTCATCTTGGCCACACCTGCTCGAATAAGTCCAGCGCGTC	1425
Db	1333	AAAGATTAATGAATTAATTCATCTTGGCCACACCTGCTCGAATAAGTCCAGCGCGTC	1392
QY	1426	AACGACTTGTTAGCCAAAACCACTCAGCTGTGATATTACATCACGTTTAAATCA	1485
Db	1393	AACGACTTGTTAGCCAAAACCACTCAGCTGTGATATTACATCACGTTTAAATCA	1452
QY	1486	GCTATTGAAGCCTGACCGCTTCAATTCAGAAATATGATTCAGTATGCAACGCTGGCA	1545
Db	1453	GCTATTGAAGCCTGACCGCTTCAATTCAGAAATATGATTCAGTATGCAACGCTGGCA	1512
QY	1546	GATGACAGCTCTGTAAATGA	1566
Db	1513	GATGACAGCTCTGTAAATGA	1533
RESULT	15		
AF102990/c			
LOCUS	AF102990	69673 bp	DNA circular BCT 18-MAY-1999
DEFINITION	Yersinia enterocolitica plasmid pYve227, complete sequence		
ACCESSION	AF102990	AF054978	AF054979
VERSION	AF054978	AF054978	AF054979
KEYWORDS	M27801	U08222	U18539
SOURCE	AF102990.1	GI:4324323	
ORGANISM	Yersinia enterocolitica		
	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		

REFERENCE
AUTHORS
TITLE
1 (bases 32056 to 32871)
Cornelis, G., Sluiter, C., de Rouvrol, C.L. and Michiels, T.
Homology between virF, the transcriptional activator of the
Yersinia virulence regulon, and ArcA, the Escherichia coli
arabinose operon regulator
J. Bacteriol. 171 (1), 254-262 (1989)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
2 (bases 15074 to 15994)
Hakansson, S., Bergman, T., Vanooteghem, J.C., Cornelis, G. and
Wolf-Watz, H.
YopB and YopD constitute a novel class of Yersinia Yop proteins
Intect. Immun. 61 (1), 71-80 (1993)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
3 (bases 45325 to 45717)
Mattiou, P. and Cornelis, G.R.
SYE, a chaperone-like protein of Yersinia enterocolitica involved
in the secretion of YopE
Mol. Microbiol. 8 (1), 123-131 (1993)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
4 (bases 23950 to 25269)
Woestyn, S., Allaoui, A., Mattiau, P. and Cornelis, G.R.
YscN, the putative energizer of the Yersinia Yop secretion
machinery
J. Bacteriol. 176 (6), 1561-1569 (1994)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
5 (bases 29897 to 30961)
Allaoui, A., Woestyn, S., Sluiter, C. and Cornelis, G.R.
YscU, a Yersinia enterocolitica inner membrane protein involved in
Yop secretion
J. Bacteriol. 176 (15), 4534-4542 (1994)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
6 (bases 43292 to 43717)
Mattiau, P., Bernier, B., Deslee, P., Michiels, T. and Cornelis, G.R.
Individual chaperones required for Yop secretion by Yersinia
Proc. Natl. Acad. Sci. U.S.A. 91 (22), 10493-10497 (1994)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
7 (bases 31537 to 31932)
Allaoui, A., Scheen, R., Lambert de Rouvrol, C. and Cornelis, G.R.
VirG, a Yersinia enterocolitica lipoprotein involved in Ca²⁺
dependency, is related to exsB of Pseudomonas aeruginosa
J. Bacteriol. 177 (15), 4230-4237 (1995)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
8 (bases 53135 to 53497)
Stainier, I., Iriarte, M. and Cornelis, G.R.
YscM and YscM2, two Yersinia enterocolitica proteins causing
downregulation of Yop transcription
Mol. Microbiol. 26 (4), 833-843 (1997)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
9 (bases 18681 to 18968)
Boyd, A.P., Sory, M.P., Iriarte, M. and Cornelis, G.R.
Heparin interferes with translocation of Yop proteins into HeLa
cells and binds to LcrE, a regulatory component of the Yersinia Yop
apparatus
Mol. Microbiol. 27 (2), 425-436 (1998)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
10 (bases 22612 to 22890)
Iriarte, M., Sory, M.P., Boland, A., Boyd, A.P., Mills, S.D.,
Lambert, I. and Cornelis, G.R.
TyeA, a protein involved in control of Yop release and in
translocation of Yersinia Yop effectors
EMBO J. 17 (7), 1907-1918 (1998)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
11 (bases 1 to 69673)
Iriarte, M., Lambert, I., Kerbourc, C. and Cornelis, G.R.
Detailed genetic map of the pYve227 plasmid of Yersinia
enterocolitica serotype O:9
Unpublished

JOURNAL
REFERENCE
AUTHORS
TITLE
12 (bases 1 to 69673)
Iriarte, M., Lambert, I., Kerbourc, C. and Cornelis, G.R.
Direct Submission
Submitted (30-Oct-1998) Microbial Pathogenesis Unit, Christian de
Dure Institute of Cellular Pathology and Faculté de Médecine.

COMMENT

Université Catholique de Louvain, Av. Hippocrate 74.49, Brussels
1200, Belgium
On May 18, 1999 this sequence version replaced g1:3411136
g1:3411135 g1:3411139 g1:3411137 g1:3603360 g1:1524380 g1:155536
g1:3603357 g1:2738986 g1:2944089 g1:3411138 g1:3108216 g1:437201
g1:497215 g1:885971 g1:155526 g1:483322 g1:48593 g1:2735421.

FEATURES

source

/organism="Yersinia enterocolitica"
/strain="W22703"
/serotype="O:9"
/db_xref="taxon:630"
/plasmid="pYec227"
33. .203
complement(889. .1434)
/note="ORF181"
/transl_table=1
/codon_start=1
/product="unknown"
/protein_id="AADI6801.1"
/db_xref="GI:4324324"
/translation="MRVVIQGGYGFKKKGLDLSMPAIOTCLGLYALSDQHDYLLCA
HDSAFGLQONLQDIFKRIEINGIRMSLRATVFGDGOSYRCSAPSSHGENEIVN
PKSHGALAOYSNQYGVPEPFNFHYOKGCGITGONRDPFLGSSPQAMLAORI
KRPSEYSLTHAMLDISKLY"
1616. .2083
/note="ORF155"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AADI6802.1"
/db_xref="GI:4324325"
/translation="MINTTELLQRIASHFGDLKLSODEYGLCELILNDRVYIMLRA
DELINLTLGLPIFGSGEARSTASQLFCYSINLNDGPGFANSEELGLIAFHL
SGELNENVSRIANFYDWLSLVSIPAEIQCLPHTQSTQSTQSVKMG"
2091. .4280
/gene="yopQ"
/note="yopQ"
2091. .4280
/gene="yopP"
/codon_start=1
/transl_table=1
/product="protein kinase yopQ"
/protein_id="AADI6803.1"
/db_xref="GI:4324326"
/translation="MKMGTPSPISLAHAHERISQWQNPVGLNIGKRRYRIIDNO
VILNPHSGFSLREGVGRKIFSGKMFNSIARLITLHAQKTTQSELRSDIPNLS
NLEGAPOTELPFGKRGKPLSGAPDLGKRAVETDKFAGESEIISIEIKNDQRLVAK
TERSIHGLFALELAEVAKHXYTAGKHPMLNANVHGAAYVPGYKREALLIMEDVGR
CSDTEIADSMKQKINSEAYVGTIKFIARHLDDTNNHAKAGIYHNDIKRQNVYFD
RSGEPVVIDGLHRSGEOPKFTSEFAPELVGNLHSEKSDVFLVYVSTLHGE
GFEKDEIRPNGLRFTISEPAHVDENGPIHRPIAGVATYTFEITIDILSDVS
RDPSENARLHEFLSDGTIDESAQIILKOTLIGEMGPLSDVATYRIPKRLRELSDLR
THLSAATKOLDWGVILSDITMLVTLDRAERGGVDKOLKFSNLILKTYVIDY
VGRREGDTSSSAEVSAPYHRSNFMLSIAEPIORIQKHLDQTFSDIGSIVAHKHL
ETILEVYVLSPOGQPVSESTYSLRLAKAKYTLSCQDLTLOQOESAKQOISILIN
RSGSADVAKQSLQRFDSIRPVKFTETQYTAHROMMAHAHAITLQEVSEFTDMRN
FTADSIPLILRIHSSSLIDELHVEOREKRELTITAIRLRLEREM"

CDS

4485. .4760
/note="ORF91A: unknown function"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AADI6804.1"
/db_xref="GI:4324327"
/translation="MLSIKRNCEPCWTLNNGLPVIGSGELAINEGKFGVNNYSALYR
IMNLGVLVRLPFCSGFTHORLNLVLQGIYVPEYKIKHTAVNSLNKRC"
4761. .5627
/gene="yopP"
/note="yopP"
4761. .5627
/gene="yopP"
/codon_start=1
/transl_table=1
/product="top effector yopP"

gene
CDSgene
CDS

/protein_id="AADI6805.1"
/db_xref="GI:4324328"
/translation="MIGPISQINSFGLSSEKTRSLISNELKNIIIOLEMDIADGSW
FKNRSLDIEVMPALVIOANKRYEMNLEFVSODSIEIKNYIENGVSRRITIN
MEGGIHREVIDYKHINKSTLLEPAPNPMGSAITLISKRTLEHYOLPDCHEFM
VEMDIOSSSEGGISLAKRLTYERSOLKITHEDNKGISDSBNLPKHLDPYL
PVTFKHQGRKRLNEYLNTNPOGQIVVKNKNETIFENPNKNSIIDGKELSVHK
KRIAEYKTLKLY"
complement(6151. .6900)
/gene="y1pA"
complement(6151. .6900)
/gene="y1pA"
/codon_start=1
/transl_table=1
/product="lipoprotein y1pA"
/protein_id="AADI6806.1"
/db_xref="GI:4324329"
/translation="MEDMKRKMLIATNAVLSVLYSGGAMSTAIKRLNLEKTKQ
MSETIWLPSQKTYLQIKNTSDKNMGLAKITKAVQDKYTVTSSPEDAHVIOA
NVLKADKMDLREAEGLFSGYOGAALGAALGIVNAGSLAGLAAGLVN
ADANVEDIYTVTDVQISEKTDPLQTDNVALKQGTSGYKVTSTQTNKHQYOTR
VYSSANKVRLKFEKPAOPVLEQDLAKSIANIL"
complement(7370. .7918)
/gene="yopQ"
complement(7370. .7918)
/gene="yopQ"
complement(7370. .7918)
/gene="yopQ"
/codon_start=1
/transl_table=1
/product="yopQ"
/protein_id="AADI6807.1"
/db_xref="GI:4324330"
/translation="MFKDAYNMRLCTALQESADDTIINTSKENNSYCATHLR
TVCYSVNVNGLIEPKSGSIIISTLELQAVQIIRLEMQVSDIDNFKLPPNSD
FELVSVLDCDGIYVTKGKDSKNGEILEYDPKNSLLENDIDIKKLYDENFRFCIML
IISSELEFSELSQDCKIMG"
8439. .9407
/gene="yopT"
8439. .9407
/gene="yopT"
/note="cytotoxin"
/codon_start=1
/transl_table=1
/product="top effector yopT"
/protein_id="AADI6808.1"

gene
CDS

Query Match 59.2%; Score 927.4; DB 2; Length 69673;
Best local similarity 97.3%; Pred. No. 7.2e-191;
Matches 943; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
586 ATGATTAGACCTAGCAACAAACCAACACATTTATTGAGATCTAGAAAAAGTTAGG 645
18679 ATGATTAGACCTAGCAACAAACCAACACATTTATTGAGATCTAGAAAAAGTTAGG 18620
646 GTGGAACAATTACTGCTATGTTCTTCAAGTTTGAAGAATTGGTTCAGTACGCAAA 705
18619 GTGGAACAATTACTGCTATGTTCTTCAAGTTTGAAGAATTGGTTCAGTACGCAAA 18560
706 GATAAAAATATGATATTTCCATTAAATATGATCCAGAAAAAGATTCGGAGTTTGGC 765
18559 GATAAAAATATGATATTTCCATTAAATATGATCCAGAAAAAGATTCGGAGTTTGGC 18500
766 AATAGATAATTACTGATGATGATGATGCTGCAAGAAAAATCCTAGCTTATTTCTACCC 825
18499 GATAGATAATTACTGATGATGATGATGCTGCAAGAAAAATCCTAGCTTATTTCTACCC 18440
826 GAGGATACCATTTCTTAAAGCGGTCTATGACACCACTGCAAAAATGGCATCAAGCGA 885
18439 GAGGATACCATTTCTTAAAGCGGTCTATGACACCACTGCAAAAATGGCATCAAGCGA 18380
886 GTAAGAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 945
18379 GTAAGAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18320

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 14:59:32 ; Search time 66.74 Seconds

(Without alignments)
5870.554 Million cell updates/sec

Title: US-08-699-716a-1
Perfect score: 1566
Sequence: 1 ATGGCCATCATCATCATCA.....ATGACACGTGTGTGAATCA 1566

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapept 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1472	94.0	1530	1	T38249
2	1409	90.0	1462	1	T38256
3	976.8	62.4	1014	1	T04222
4	975.2	62.3	1014	1	T38242
5	974.8	62.2	1014	1	T38243
6	973.2	62.1	1014	1	T04223
7	517.2	33.0	544	1	V41594
8	516.8	33.0	544	1	V41596
9	516.8	33.0	544	1	T38248
10	510	32.6	510	1	V41595
11	452.8	28.9	474	1	V41600
12	452.2	28.9	541	1	O92817
13	452.2	28.9	541	1	T38244
14	451	28.8	542	1	O92818
15	448	28.6	450	1	V41601
16	447	28.5	447	1	V41609
17	402	25.7	576	1	V41598
18	339	21.6	513	1	V41599
19	68.8	4.4	861	1	O81500
20	68.4	4.4	861	1	T45347
21	68.4	4.4	375	1	V54056
22	68.4	4.4	375	1	V54057
23	68.4	4.4	375	1	V54058
24	68.4	4.4	375	1	V54059
25	54	3.4	1546	1	V30575
26	53.4	3.4	1278	1	V41263
27	53.4	3.4	1402	1	T29246
28	53.4	3.4	1402	1	V30572
29	53.4	3.4	1460	1	V30583
30	53.4	3.4	1463	1	V30584
31	53.4	3.4	1469	1	V30591
32	53.4	3.4	1472	1	V30585
33	53.4	3.4	1502	1	V30588

34	53.4	3.4	1526	1	V30580	Clostridium botuli
35	53.4	3.4	1535	1	V30596	Clostridium botuli
36	53.4	3.4	1547	1	V30581	Clostridium botuli
37	46.8	3.0	104	1	T33359	Ich-2 gene 5' prim
38	45	2.9	21591	1	X13047	Enterococcus faeca
39	42	2.7	110000	1	X20248_01	Continuation (2 of
40	42	2.7	110000	1	X20248_02	Continuation (3 of
41	40	2.6	110000	1	V21309_15	Continuation (16 of
42	39	2.5	4435	1	O29686	CCV-c54 spike gene
43	38.8	2.5	1116	1	O52273	Soybean seed stora
44	38.6	2.5	1124	1	T29508	CONSTANS gene. Iso
45	38	2.4	7953	1	V74463	Staphylococcus aur

ALIGNMENTS

RESULT 1
T38249
ID T38249 standard; DNA; 1530 BP.
AC T38249:1996 (first entry)
DE Y. pestis FI/V antigen gene fusion.
KW Plague; vaccine; genetic immunisation; V antigen; lcrv;
KW FI antigen; cafl; ds.
OS Chimeric Yersinia pestis strain GB;
FH Key
FT cds
FT Location/Qualifiers
FT /tag- a
FT /product- FI/V fusion protein
FT signal_peptide 13..75
FT /tag- b
FT mat_peptide 76..1515
FT /tag- c
FT mutation 1522
FT /tag- d
FT /note- "base T at position 1522 was modified
from G to create a second in-frame stop
codon"

FT WO9628551-A1.
FT 19-SEP-1996.
FT 13-MAR-1996: G00571.
FT 13-MAR-1995: GB-005059.
FT 15-SEP-1995: GB-018946.
FT 05-DEC-1995: GB-024825.
FT PA (MINA) UK SEC FOR DEFENCE.
FT PI Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
FT WPI: 96-433824/43.
FT P-PSDS: W01044.
FT Yersinia pestis V antigen and FI antigen or their protective
FT PT epitopic parts - useful in vaccine for protection against plague
FT PS Claim 41; Page 65-69; 98pp; English.
FT A gene fusion (T38249) comprises coding sequences for the Yersinia
FT CC pestis FI antigen (see also T38248) including the signal sequence
FT CC and for the Y. pestis V antigen (see also T38243), joined by a
FT CC sequence encoding a 6-amino acid peptide linker. It was obtd. by
FT CC PCR amplification of the individual genes using primers (see also
FT CC T38245) based on the genes and including linker sequences. The
FT CC gene fusion (see also T38256) can be used to produce FI/V fusion
FT CC protein (W01044) in transformed cells, esp. gut-colonising, the
FT CC organisms, to induce an immune response against Y. pestis, the
FT CC causative organism of plague.
FT SO Sequence 1530 BP; 499 A; 296 C; 312 G; 423 T;

Query Match 94.0%; Score 1472; DB 1; Length 1530;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 149; Conservative 0; Mismatches 5; Indels 9; Gaps 1;
OY 68 ATATGAAAAATCATGTCCTATCCCATATGATATTTGGAACATGCAACTGCTA 127
DB 11 ATATGAAAAATCATGTCCTATCCCATATGATATTTGGAACATGCAACTGCTA 70

OY	128	ATGGGCGATTTAACAGCAAGCAACCACTGCACGGCACTCTGTGTAAACAGCCGCA	187
Db	71	ATGGCGCGATTTAACATGCAAGCAACCACTGCACGGCACTCTGTGTAAACAGCCGCA	130
OY	188	TCACCTTACATATAGAAAGCGCTCCATTCATTAATGACAAATGAAACATCGATA	247
Db	131	TCACCTTACATATAGAAAGCGCTCCATTCATTAATGACAAATGAAACATCGATA	190
OY	248	CAGAACTACTGTTGGTACGCTTAACCTTGGCGGCTATATAACAGAACCACTACACAT	307
Db	191	CAGAACTACTGTTGGTACGCTTAACCTTGGCGGCTATATAACAGAACCACTACACAT	250
OY	308	CTGTAACTTACAGATGCGCGGGGTATCCATGTACTTAACATTAATTAATCTCAGAGATG	367
Db	251	CTGTAACTTACAGATGCGCGGGGTATCCATGTACTTAACATTAATTAATCTCAGAGATG	310
OY	368	GAAATTAACCCAAATTCATCTACAAAAGTAGTTGGCAGAGATTCAGAGATTTTGATATCT	427
Db	311	GAAATTAACCCAAATTCATCTACAAAAGTAGTTGGCAGAGATTCAGAGATTTTGATATCT	370
OY	428	CTCCTAAGGTAAAGGTGAGACCTTGTTGGGGATACGCGCTGGCTACGGCAGCC	487
Db	371	CTCCTAAGGTAAAGGTGAGACCTTGTTGGGGATACGCGCTGGCTACGGCAGCC	430
OY	488	AGGATTTCTTTGTCGCTCAATTGGTCCAAAGCGGTAAACTTGACAGCAGGTAATAACA	547
Db	431	AGGATTTCTTTGTCGCTCAATTGGTCCAAAGCGGTAAACTTGACAGCAGGTAATAACA	490
OY	548	CTGATGCGTAAACGTTAAACCGTATCTAACCAAAATTCAT-----GATTAAAGCT	598
Db	491	CTGATGCGTAAACGTTAAACCGTATCTAACCAAAAGTTCATGATTAAGAGCT	550
OY	599	ACGAACAAAACCCCAACATTTTATGGAGATCTAGAAAAAGTAGAGGTGGAACAACCTTA	658
Db	551	ACGAACAAAACCCCAACATTTTATGGAGATCTAGAAAAAGTAGAGGTGGAACAACCTTA	610
OY	659	CTGGTCATGGTCTTCCAGTTTAGAGAAATGGTTCAGTTAGTCAAGATATAAATAATAG	718
Db	611	CTGGTCATGGTCTTCCAGTTTAGAGAAATGGTTCAGTTAGTCAAGATATAAATAATAG	670
OY	719	ATATTCCATTAATTAATGATCCCGAAGAAATCGGAGGTTTTGGCAATAGAGTAATA	778
Db	671	ATATTCCATTAATTAATGATCCCGAAGAAATCGGAGGTTTTGGCAATAGAGTAATA	730
OY	779	CTGATGATTCGAATTCCTCAAGAAAATCCTAGCTTAATTTCTACCCGAGATACCATTC	838
Db	731	CTGATGATTCGAATTCCTCAAGAAAATCCTAGCTTAATTTCTACCCGAGATACCATTC	790
OY	839	TTAAAGGGGGCATTAAGAACCAACCTGCAAAATGGCATCAAGGAGTAAGAAAGATCC	898
Db	791	TTAAAGGGGGCATTAAGAACCAACCTGCAAAATGGCATCAAGGAGTAAGAAAGATCC	850
OY	899	TTGATCATCGCCGAATACAAATGGGAATGCGGGCGTCAATGGCAGTAATGATTTCT	958
Db	851	TTGATCATCGCCGAATACAAATGGGAATGCGGGCGTCAATGGCAGTAATGATTTCT	910
OY	959	CTTTAACCCGCGATCGTATGATGATGATATTTTGAAGTAGTTGTGATTCATGATGATC	1018
Db	911	CTTTAACCCGCGCGATCGTATGATGATGATATTTTGAAGTAGTTGTGATTCATGATGATC	970
OY	1019	ATCATGGTATGCGCCGTAAGCAAGTGGGTGAAGAAATTAATCTGAGGCTTACCGCCGAATTA	1078
Db	971	ATCATGGTATGCGCCGTAAGCAAGTGGGTGAAGAAATTAATCTGAGGCTTACCGCCGAATTA	1030
OY	1079	AGATTTATTCAGTTATTCAGCGGAAATTAATAGACATCTGCTAGTAGTGCCACCATTA	1138
Db	1031	AGATTTATTCAGTTATTCAGCGGAAATTAATAGACATCTGCTAGTAGTGCCACCATTA	1090
OY	1139	ATATCCCAAGTAATTCATTAATCTCATGATATAAATTAATATGTTATACAGATGAG	1198
Db	1091	ATATCCCAAGTAATTCATTAATCTCATGATATAAATTAATATGTTATACAGATGAG	1150

OY	1159	AGATTTTAAAGCCAGGCGAGAGTACAAAATTCTCGAAGAAATGCTCAACACCACTTC	1258
OY	1159	AGATTTTAAAGCCAGGCGAGAGTACAAAATTCTCGAAGAAATGCTCAACACCACTTC	1258
DB	1151	AGATTTTAAAGCCAGGCGAGAGTACAAAATTCTCGAAGAAATGCTCAACACCACTTC	1210
OY	1259	AGGTGATGGAGCGAGAAAAAATAGTCTCGATAAAGACCTTCTTGAAAGTGAATA	1318
DB	1211	AGGTGATGGAGCGAGAAAAAATAGTCTCGATAAAGACCTTCTTGAAAGTGAATA	1270
OY	1319	AAAGACCGGGGGCTTGCGTATCTGATAAAACGACACCTTATATAAGTATATATG	1378
DB	1271	AAAGACCGGGGGCTTGCGTATCTGATAAAACGACACCTTATATAAGTATATATG	1350
OY	1379	AATATCTCACTTTGGCCACCACTGCTCGGATTAAGTCAGGCGCTCAACGACTTGTTA	1438
DB	1331	AATATCTCACTTTGGCCACCACTGCTCGGATTAAGTCAGGCGCTCAACGACTTGTTA	1390
OY	1439	GCCAAAAAACACACTCAGCTCTCTGATATTACATCACGTTTATTCAGCTATTGAAGAC	1498
DB	1391	GCCAAAAAACACACTCAGCTCTCTGATATTACATCACGTTTATTCAGCTATTGAAGAC	1450
OY	1499	TGAACGCTTCATTCAGAAATATGATTCAGTGATGCAACGCTGCTAGATGACAGCTG	1558
DB	1451	TGAACGCTTCATTCAGAAATATGATTCAGTGATGCAACGCTGCTAGATGACAGCTG	1510
OY	1559	GTAATGA 1566	
DB	1511	GTAATGA 1518	
RESULT 2			
PD	T38256		
AC	T38256	standard: DNA; 1462 BP.	
DT	28-DEC-1996	(first entry)	
DE	Y. pestis Fl/V antigen gene fusion.		
KW	Plague: vaccine; genetic immunisation: V antigen; lcrV;		
KM	Fl antigen; calf; ds.		
OS	Chimeric Yersinia pestis strain GB.		
OS	Chimeric synthetic.		
FT	Key	Location/Qualifiers	
FT	cds	8..1450	
FT		/*tag- a	
FT		/product- Fl/V fusion protein	
FT		452..472	
FT		/*tag- b	
FT	misc.feature	/note- "bases 452-472 is a sequence derived from PCR primers"	
PD	MO9628551-A1.		
PD	19-SEP-1996		
PR	13-MAR-1996;	G00571.	
PR	13-MAR-1995;	GB-005059.	
PR	15-SEP-1995;	GB-018946.	
PR	05-DEC-1995;	GB-024825.	
PA	(MINA) UK SEC FOR DEPENDENCE.		
PI	Bennett AM, Leary SEC, Oyston PCF, Tiltball RM, Williamson ED;		
DR	WPI: 96-433824/43.		
DR	P-RSDS; W01045.		
PT	Yersinia pestis V antigen and Fl antigen or their protective		
PT	epitopic parts - useful in vaccine for protection against plague		
PS	Disclosure; Page 51-55; 98pb; English.		
CC	A gene fusion (T38256) comprises coding sequences for the Yersinia		
CC	pestis Fl antigen (see also T38244) (without the signal sequence)		
CC	and for the Y pestis V antigen (see also T38243). joined by a		
CC	sequence encoding a 6-amino acid peptide linker. It was obd. by		
CC	PCR amplification of the individual genes using primers (see also		
CC	T38245) based on the genes and including linker sequences. The		
CC	gene fusion (see also T38249) can be used to produce Fl/V fusion		
CC	protein (W01045) in transformed cells, esp. gut-colonising		
CC	organisms, to induce an immune response against Y. pestis, the		
CC	causative organism of plague.		
SO	Sequence 1462 BP; 4/6 A; 285 C; 301 G; 400 T;		


```

Db 11 TCATTAGAGCCTAGCAACAAACCCACACATTTTATTAGAGATCTAATAAAAGTTAGCG 70
QY 647 TGAACAACACTACTGTCATGCTGTTCTTCAAGTTTGAAGAATGGTTCAGTTAGTCAAG 706
Db 71 TGAACAACACTACTGTCATGCTGTTCTTCAAGTTTGAAGAATGGTTCAGTTAGTCAAG 130
QY 707 ATAAAAATATAGATATTTCCATTAATATGATCCAGAAAAGATTCGAGGTTTGGCA 766
Db 131 ATAAAAATATAGATATTTCCATTAATATGATCCAGAAAAGATTCGAGGTTTGGCA 190
QY 767 ATAGAGATATCTAGATGATATCCGATTCGACAGAAAATCTTACGTTATTTTACCCG 826
Db 191 ATAGAGATATCTAGATGATATCCGATTCGACAGAAAATCTTACGTTATTTTACCCG 250
QY 827 AGATATACATCTTAAAGGCGGTATTATGACAAACCACTGCAAAATGGCATCAAGCGAG 886
Db 251 AGATATACATCTTAAAGGCGGTATTATGACAAACCACTGCAAAATGGCATCAAGCGAG 310
QY 887 TAAAGAGTTCCTTGAATCATCGCGGAATACACAATGGGAATTCGCGGCGTTTCATGGCAG 946
Db 311 TAAAGAGTTCCTTGAATCATCGCGGAATACACAATGGGAATTCGCGGCGTTTCATGGCAG 370
QY 947 TAATGATTTCTCTTAAACCCCATGCTATGATGATGATTTTGAAGTGTGTTG 1006
Db 371 TAATGATTTCTCTTAAACCCCATGCTATGATGATGATTTTGAAGTGTGTTG 430
QY 1007 ATTAATGATATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
Db 431 ATTAATGATATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 490
QY 1067 CCGCGCAATTAAGATTTTCACTTATTCAGCGCAATTAATTAAGCATCTGTCTAGTA 1126
Db 491 CCGCGCAATTAAGATTTTCACTTATTCAGCGCAATTAATTAAGCATCTGTCTAGTA 550
QY 1127 GTGCGACCATTAATATCCATATTAATCCATTAATCTCATGATGATGATGATGATGATGAT 1186
Db 551 GTGCGACCATTAATATCCATATTAATCCATTAATCTCATGATGATGATGATGATGATGAT 610
QY 1187 ATACAGATGAAGAGATTTTAAAGCCAGCGAGATCAAAATTTCTCGAAGAAATGCTC 1246
Db 611 ATACAGATGAAGAGATTTTAAAGCCAGCGAGATCAAAATTTCTCGAAGAAATGCTC 670
QY 1247 AAACCCACATTCAGGTGATGAGCGAGCAAGAAAATAGTCTCGATTAAGGACTTTCTG 1306
Db 671 AAACCCACATTCAGGTGATGAGCGAGCAAGAAAATAGTCTCGATTAAGGACTTTCTG 730
QY 1307 GAATGAGATTAAGAAAGAACCGGGGCGTTGGTATCTGAAAACTCATCTCTATATA 1386
Db 731 GAATGAGATTAAGAAAGAACCGGGGCGTTGGTATCTGAAAACTCATCTCTATATA 790
QY 1367 AAGATTAATTAAGATTAATCTCACTTTGGCACCACTGCTGGATTAAGCCAGGCGGCTCA 1426
Db 791 AAGATTAATTAAGATTAATCTCACTTTGGCACCACTGCTGGATTAAGCCAGGCGGCTCA 850
QY 1427 ACAGCTGGTTAGCCAAACCAACTCAGCTGTCTGATTAATTAATCAAGTTTAAATCAG 1486
Db 851 ACAGCTGGTTAGCCAAACCAACTCAGCTGTCTGATTAATTAATCAAGTTTAAATCAG 910
QY 1487 CTATTGAAGCCTGAACCGTTTCATTCAGAAATATGATTCGATGATGATGATGATGATG 1546
Db 911 CTATTGAAGCCTGAACCGTTTCATTCAGAAATATGATTCGATGATGATGATGATGATG 970
QY 1547 ATGACAGCTCTGTAATGA 1566
Db 971 ATGACAGCTCTGTAATGA 990

```

```

DE Y. pestis lcrV (V antigen) gene.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW Fl antigen; ds.
OS Yersinia pestis
FH Key Location/Qualifiers
FT cds 1..990
FT misc_feature 1..6
FT FT /*tag- a
FT FT /*tag- b
FT FT /note- "vector pMAL-p2 or pMAL-c2-derived bases"
FT FT mutation 1006
FT FT /*tag- c
FT FT /note- "base 1006 is altered to a T to create a
FT FT second in-frame stop codon"
MO9628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996: G00571.
PR 13-MAR-1995: GB-005059.
PR 15-SEP-1995: GB-018946.
PR 05-DEC-1995: GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tibball RM, Williamson ED,
DR WPI; 96-433824/43.
DR P-PSDB; W01040.
PT Yersinia pestis V antigen and Fl antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Claim 5; Page 25-28; 98pp; English.
CC A lcrV gene sequence (T38242) codes for the Yersinia pestis V
CC antigen (W01040), which is capable of evoking protective immune
CC responses in animals. The gene was amplified from Y. pestis
CC DNA by PCR using primers (T38250-51) homologous to the 5' and 3'
CC ends of the gene. The gene was inserted into vector pMAL-p2,
CC pMAL-c2 or pGEX-5x-2 (see also T38243) to allow prodn. of
CC recombinant V antigen for use in vaccines against plague.
CC Expression in gut-colonising organisms and attenuated Salmonella
CC typhl allows prodn. of live vaccines. Fl/V antigen fusions were
CC also created (see also T38249 and T38256). The gene can itself be
CC used in genetic vaccines.
SQ Sequence 1014 bp; 347 A; 180 C; 201 G; 286 T;

Query Match 62.3%; Score 975.2; DB 1; Length 1014;
Best Local Similarity 99.7%; Pred. No. 1.9e-250;
Matches 977; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

RESULT 4
 T38242
 ID T38242 standard: DNA; 1014 bp.
 AC T38242;
 DT 28-DEC-1996 (first entry)

QY	1007	ATTCAAATGAATCAATGAGGATGCGCGGTGACAGTGGGTGAAGATTAAGTGAAGCTTA	1066
Db	431	ATTCAATGAATCATCATGATGATGATGCCCGTACAGATTGGCTGAGATTAAGTGAAGCTTA	490
QY	1067	CCGCGGATTAAGAAGATTTATTCAGTATTCACGCCCAATTAATAAGCATCTGTATAGTA	1126
Db	491	CCGCGGATTAAGAAGATTTATTCAGTATTCACGCCCAATTAATAAGCATCTGTATAGTA	550
QY	1127	GTGGCACCATTAAATATCCATGATTAATTCATTATATCTCATGTGATAAAAATTTATATGTT	1186
Db	551	GTGGCACCATTAAATATCCATGATTAATTCATTATATCTCATGTGATAAAAATTTATATGTT	610
QY	1187	ATACAGATGAAGAGATTTTAAAGCCAGCCGACAGTACAAATTCCTCGAGAAATGCTTC	1246
Db	611	ATACAGATGAAGAGATTTTAAAGCCAGCCGACAGTACAAATTCCTCGAGAAATGCTTC	670
QY	1247	AAACACCATTCACAGTGTGATGGAGGAGAGAAAAAATAGTCTCGATTAAGAGACTTCTTG	1306
Db	671	AAACACCATTCACAGTGTGATGGAGGAGAGAAAAAATAGTCTCGATTAAGAGACTTCTTG	730
QY	1307	GAACTGAGAAATAAAAGAACCGGGGCGTTGGGTATCTGAAAAAAGCTATACCTTTAATA	1366
Db	731	GAACTGAGAAATAAAAGAACCGGGGCGTTGGGTATCTGAAAAAAGCTATACCTTTAATA	790
QY	1367	AAGTATTAATGAATATCTCATCTTTGCCACCACTGCTCGGATTAAGTCCAGGCGCTCA	1426
Db	791	AAGTATTAATGAATATCTCATCTTTGCCACCACTGCTCGGATTAAGTCCAGGCGCTCA	850
QY	1427	ACGACTTGTTAGCCAAAAAAGCACTCAGCTGCTGATATTAATCATCAGCTTTAATTAG	1486
Db	851	ACGACTTGTTAGCCAAAAAAGCACTCAGCTGCTGATATTAATCATCAGCTTTAATTAG	910
QY	1487	CTATTGAGCACTGAACCGCTTCATTCAGAAATATGATTCAGTATGCAACGCTGCTAG	1546
Db	911	CTATTGAGCACTGAACCGCTTCATTCAGAAATATGATTCAGTATGCAACGCTGCTAG	970
QY	1547	ATGACACGCTCTGTTAATGA	1566
Db	971	ATGACACGCTCTGTTAATGA	990
RESULT	5		
T38243			
ID	T38243	standard; DNA; 1014 BP.	
AC	T38243		
DT	28-DEC-1996	(first entry)	
DE	X. pestis lcrV (V antigen) gene.		
KW	Plague; vaccine; genetic immunisation; V antigen; lcrV;		
KW	F1 antigen; ds.		
OS	Yersinia pestis strain GB.		
FH	Key	location/Qualifiers	
FT	cds	1..990	
FT		/*tag- a	
FT	misc-feature	1..10	
FT		/*tag- b	
FT	mutation	/note- "vector pGEX-5X-2-derived bases"	
FT		16	
FT		/*tag- c	
FT	mutation	/note- "base 16 is altered to a C from an A to create an EcoRI site"	
FT		1006	
FT		/*tag- d	
FT		/note- "base 1006 is altered to a T to create a second in-frame stop codon"	
PN	W09628551-A1.		
PD	19-SEP-1996.		
PF	13-MAR-1996; G00571.		
PR	15-SEP-1995; GB-005059.		
PR	15-SEP-1995; GB-018946.		
PR	05-DEC-1995; GB-024825.		
PA	(MINA) UK SEC FOR DEFENCE.		
PI	Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;		
PR	WPI; 96-433824/43.		

Query Match	Best Local Similarity	Score	DB 1	Length	1014	DB 2	Length	1014	DB 3	Length	1014	DB 4	Length	1014	DB 5	Length	1014	DB 6	Length	1014	DB 7	Length	1014	DB 8	Length	1014	DB 9	Length	1014	DB 10	Length	1014	DB 11	Length	1014	DB 12	Length	1014	DB 13	Length	1014	DB 14	Length	1014	DB 15	Length	1014	DB 16	Length	1014	DB 17	Length	1014	DB 18	Length	1014	DB 19	Length	1014	DB 20	Length	1014	DB 21	Length	1014	DB 22	Length	1014	DB 23	Length	1014	DB 24	Length	1014	DB 25	Length	1014	DB 26	Length	1014	DB 27	Length	1014	DB 28	Length	1014	DB 29	Length	1014	DB 30	Length	1014	DB 31	Length	1014	DB 32	Length	1014	DB 33	Length	1014	DB 34	Length	1014	DB 35	Length	1014	DB 36	Length	1014	DB 37	Length	1014	DB 38	Length	1014	DB 39	Length	1014	DB 40	Length	1014	DB 41	Length	1014	DB 42	Length	1014	DB 43	Length	1014	DB 44	Length	1014	DB 45	Length	1014	DB 46	Length	1014	DB 47	Length	1014	DB 48	Length	1014	DB 49	Length	1014	DB 50	Length	1014	DB 51	Length	1014	DB 52	Length	1014	DB 53	Length	1014	DB 54	Length	1014	DB 55	Length	1014	DB 56	Length	1014	DB 57	Length	1014	DB 58	Length	1014	DB 59	Length	1014	DB 60	Length	1014	DB 61	Length	1014	DB 62	Length	1014	DB 63	Length	1014	DB 64	Length	1014	DB 65	Length	1014	DB 66	Length	1014	DB 67	Length	1014	DB 68	Length	1014	DB 69	Length	1014	DB 70	Length	1014	DB 71	Length	1014	DB 72	Length	1014	DB 73	Length	1014	DB 74	Length	1014	DB 75	Length	1014	DB 76	Length	1014	DB 77	Length	1014	DB 78	Length	1014	DB 79	Length	1014	DB 80	Length	1014	DB 81	Length	1014	DB 82	Length	1014	DB 83	Length	1014	DB 84	Length	1014	DB 85	Length	1014	DB 86	Length	1014	DB 87	Length	1014	DB 88	Length	1014	DB 89	Length	1014	DB 90	Length	1014	DB 91	Length	1014	DB 92	Length	1014	DB 93	Length	1014	DB 94	Length	1014	DB 95	Length	1014	DB 96	Length	1014	DB 97	Length	1014	DB 98	Length	1014	DB 99	Length	1014	DB 100	Length	1014	DB 101	Length	1014	DB 102	Length	1014	DB 103	Length	1014	DB 104	Length	1014	DB 105	Length	1014	DB 106	Length	1014	DB 107	Length	1014	DB 108	Length	1014	DB 109	Length	1014	DB 110	Length	1014	DB 111	Length	1014	DB 112	Length	1014	DB 113	Length	1014	DB 114	Length	1014	DB 115	Length	1014	DB 116	Length	1014	DB 117	Length	1014	DB 118	Length	1014	DB 119	Length	1014	DB 120	Length	1014	DB 121	Length	1014	DB 122	Length	1014	DB 123	Length	1014	DB 124	Length	1014	DB 125	Length	1014	DB 126	Length	1014	DB 127	Length	1014	DB 128	Length	1014	DB 129	Length	1014	DB 130	Length	1014	DB 131	Length	1014	DB 132	Length	1014	DB 133	Length	1014	DB 134	Length	1014	DB 135	Length	1014	DB 136	Length	1014	DB 137	Length	1014	DB 138	Length	1014	DB 139	Length	1014	DB 140	Length	1014	DB 141	Length	1014	DB 142	Length	1014	DB 143	Length	1014	DB 144	Length	1014	DB 145	Length	1014	DB 146	Length	1014	DB 147	Length	1014	DB 148	Length	1014	DB 149	Length	101
-------------	-----------------------	-------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	-------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	------	--------	--------	-----

```

QY 1369 GATATAATGATATATCTCACTTGTGCACACCTGCTCGATAGATCCAGCCGCTCAAC 1428
DB 793 GATAATATGATATATCTCACTTGTGCACACCTGCTCGATAGATCCAGCCGCTCAAC 852
QY 1429 GACTTGGTTAGCCAAAACAACTACAGCTGTGTGATATTTACATCAGCTTTTAATTCAGCT 1488
DB 853 GACTTGGTTAGCCAAAACAACTACAGCTGTGTGATATTTACATCAGCTTTTAATTCAGCT 912
QY 1489 ATGAAGACATGACCTTTCATTCAGAAATATGATTCAGTATGACAGCTGCTAGAT 1548
DB 913 ATGAAGACATGACCTTTCATTCAGAAATATGATTCAGTATGACAGCTGCTAGAT 972
QY 1549 GACACGCTGTGTAATGA 1566
DB 973 GACACGCTGTGTAATGA 990

RESULT 6
T04223
ID T04223 standard; DNA; 1014 BP.
AC T04223;
DT 18-APR-1996 (first entry)
DE Partial LcrV (V antigen) gene of Y. pestis.
KW LcrV; V antigen; virulence; plague; vaccine; epitope; ss.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT cds 1..990
FT FT /*tag= a
FT PN MO9524475-A1.
FT PD 14-SEP-1995.
FT PF 06-MAR-1995; G00481.
FT PR 08-MAR-1994; GB-004577.
FT PA (MINA) UK SEC FOR DEFENCE.
FT PI Leary SEC, Tilball RW, Williamson ED, Leary SE;
DR MPI, 95-328268/42.
DR P-PSDB; R/79962.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PT oral or parenteral vaccines for protection against plague
PT Claim 6; Page 15-16; 25pp; English.
CC T04222-73 are DNA sequences (LcrV) encoding all or a protective epitopic
CC part of the mature V protein of Yersinia pestis. The protein was
CC expressed as a fusion protein with maltose binding protein. Y. pestis
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis
CC is the highly virulent causative organism of plague in a wide range of
CC animals, including man. The V antigen (LcrV) is an unstable 37.3 kDa
CC monomeric peptide encoded on the ca. 70 kb Lcr plasmid. The V antigen
CC is postulated to act as a virulence antigen, and transformed
CC microorganisms contg. recombinant DNA encoding a V antigen protein/
CC peptide are useful in vaccines to protect against plague.
SQ Sequence 1014 BP; 343 A; 185 C; 204 G; 282 T;

Query Match 62.1%; Score 973.2; DB 1; Length 1014;
Best Local Similarity 99.7%; Pred. No. 6; 6e-250;
Matches 975; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 829 GATACATCTTTAAAGCGGTCATTATGACAACTGCAAAATGGCATCAAGCAGTA 888
DB 253 GATGCCATCTTTAAAGCGGTCATTATGACAACTGCAAAATGGCATCAAGCAGTA 312
QY 889 AAGAGTCTCTTAATCATTCGCCGAATACACATGGGAATTTGGCGGCTTCATGGCAGTA 948
DB 313 AAGAGTCTCTTAATCATTCGCCGAATACACATGGGAATTTGGCGGCTTCATGGCAGTA 372
QY 949 ATGATCTCTTTAAAGCGGTCATTATGACAACTGCAAAATGGCATCAAGCAGTA 1008
DB 373 ATGATCTCTTTAAAGCGGTCATTATGACAACTGCAAAATGGCATCAAGCAGTA 432
QY 1009 TCAATGATCATCATGATGATGCGGTCAGCAAGTTGCGTGAAGAAATTAAGCTTAC 1068
DB 433 TCAATGATCATCATGATGATGCGGTCAGCAAGTTGCGTGAAGAAATTAAGCTTAC 492
QY 1069 GCGGATTTAAAGTATTATTCAGTATTCAGACCGCAAAATTAATGATGCTGTAGTACT 1128
DB 493 GCGGATTTAAAGTATTATTCAGTATTCAGACCGCAAAATTAATGATGCTGTAGTACT 552
QY 1129 GGCACCATTAATATTCATGATTAATCATTAATCATGATTAATTAATGATGCTGTAT 1188
DB 553 GGCACCATTAATATTCATGATTAATCATTAATCATGATTAATTAATGATGCTGTAT 612
QY 1189 ACAGATGAAGAGATTTTAAAGCGCAGCAGATACAAATTTCTGAGAAATGGCTCAA 1248
DB 613 ACAGATGAAGAGATTTTAAAGCGCAGCAGATACAAATTTCTGAGAAATGGCTCAA 672
QY 1249 ACCACCATTCAGTGTGATGAGGAGCGAGAAAATATGCTCGATTAAGAGCTTTCTGGA 1308
DB 673 ACCACCATTCAGTGTGATGAGGAGCGAGAAAATATGCTCGATTAAGAGCTTTCTGGA 732
QY 1309 AGTGAATTAAGAAAGACCGGGGCTGGTGAATCTGAAATGCAATCTTATATATAA 1368
DB 733 AGTGAATTAAGAAAGACCGGGGCTGGTGAATCTGAAATGCAATCTTATATATAA 792
QY 1369 GATATAATGATATATCTCACTTGTGCACACCTGCTCGATAGTCCAGCCGCTCAAC 1428
DB 793 GATAATATGATATATCTCACTTGTGCACACCTGCTCGATAGTCCAGCCGCTCAAC 852
QY 1429 GACTTGGTTAGCCAAAACAACTACAGCTGTGTGATATTTACATCAGCTTTTAATTCAGCT 1488
DB 853 GACTTGGTTAGCCAAAACAACTACAGCTGTGTGATATTTACATCAGCTTTTAATTCAGCT 912
QY 1489 ATGAAGACATGACCTTTCATTCAGAAATATGATTCAGTATGACAGCTGCTAGAT 1548
DB 913 ATGAAGACATGACCTTTCATTCAGAAATATGATTCAGTATGACAGCTGCTAGAT 972
QY 1549 GACACGCTGTGTAATGA 1566
DB 973 GACACGCTGTGTAATGA 990

RESULT 7
V41594
ID V41594 standard; DNA; 544 BP.
AC V41594;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of FI antigen nYPI(a)sec544.
KW FI antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT cds 17..529
FT FT /*tag= a
FT PN "FI antigen"
FT PD WO9824912-A2.
FT PF 11-JUN-1998.
FT PR 04-DEC-1997; U22617.
FT PA (HESK-) HESKA CORP.
FT PI Haines EJ, Osorio JE, Thomas RE;
DR MPI, 98-333331/29.
DR P-PSDB; W59782.

```

PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Page 51; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 544 BP; 164 A; 120 C; 114 G; 146 T;

Query Match 33.0%; Score 517.2; DB 1; Length 544;
 Best Local Similarity 98.5%; Pred. No. 1.4e-128;
 Matches 522; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATACGTCGTTATCCGATGATATTTGGAACTATTCGAACTGCTA 127
 DB 15 ATATGAAAAAATACGTCGTTATCCGATGATATTTGGAACTATTCGAACTGCTA 74
 QY 128 ATGGGAGATTATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 187
 DB 75 ATGGGAGATTATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 134
 QY 188 TCACTCTTACATATATAGAAAGCGCTCCATTTACATATATGCAATGAAACATGATA 247
 DB 135 TCACTCTTACATATATAGAAAGCGCTCCATTTACATATATGCAATGAAACATGATA 194
 QY 248 CAGATTAAGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 307
 DB 195 CAGATTAAGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 254
 QY 308 CTGTAACTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 367
 DB 255 CTGTAACTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 314
 QY 368 GAAATTAACCAACCAATTAACCAACCAATTAACCAACCAATTAACCAATTAACCA 427
 DB 315 GAAATTAACCAACCAATTAACCAACCAATTAACCAACCAATTAACCAATTAACCA 374
 QY 428 CTCTTAAGTAAAGGAG 487
 DB 375 CTCTTAAGTAAAGGAG 434
 QY 488 AGGATTTCTTTGTCGTCATATGTTCCAAAGCGGTAACCTTGCAGCAGTAATACA 547
 DB 435 AGGATTTCTTTGTCGTCATATGTTCCAAAGCGGTAACCTTGCAGCAGTAATACA 494
 QY 548 CTGATGCTGTAACCGTAACCGTAATCTAACAAGATTCATGATTAGAGCC 597
 DB 495 CTGATGCTGTAACCGTAACCGTAATCTAACAAGATTCATGATTAGAGATCC 544

RESULT 8
 ID 092819 standard; DNA; 542 BP.
 AC 092819;
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis cafi (F1) antigen in plasmid pFORF1b.
 KW Yersinia pestis; Yersinia pestis; Yersinia pestis; Yersinia pestis;
 KM Yersinia pestis; Yersinia pestis; Yersinia pestis; Yersinia pestis;
 OS Yersinia pestis.
 FH Key
 FT cds
 FT 2-7
 FT /tag- a
 FT /note- "first protein encoded by pFORF1b"
 FT misc-feature
 FT 1.6
 FT /tag- b
 FT misc-feature
 FT 536..541
 FT /tag- c
 FT /note- "cafi open reading frame downstream seq."
 FT 21..530
 FT /tag- d
 FT /note- "cafi fusion protein"

PN WC9518231-A1.
 PD 06-JUL-1995.
 PE 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tildall RW, Williamson ED;
 DR P-SDS; R76528.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure; Page 19-20; 27pp; English.
 CC The sequence represents the plasmid pFORF1b including the entire
 CC Y. pestis cafi (F1) antigen gene having a 5' tail including a SacI
 CC restriction site, and up to 1400 bp downstream of the cafi ORF.
 CC The DNA construct can be used to transform human or animal gut
 CC colonizing microorganisms, specifically attenuated salmonella
 CC typhimurium or Yersinia pestis. The transformed microorganisms
 CC can be used as live/attenuated vaccines which induce immune
 CC responses at mucosal surfaces. The vaccines provide protection
 CC against infection with Y. pestis, and are parenterally and orally
 CC active vaccines offering protection against bubonic and pneumonic
 CC plague.
 SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

Query Match 33.0%; Score 516.8; DB 1; Length 542;
 Best Local Similarity 99.0%; Pred. No. 1.8e-128;
 Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATACGTCGTTATCCGATGATATTTGGAACTATTCGAACTGCTA 127
 DB 19 ATATGAAAAAATACGTCGTTATCCGATGATATTTGGAACTATTCGAACTGCTA 78
 QY 128 ATGGGAGATTATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 187
 DB 79 ATGGGAGATTATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 138
 QY 188 TCACTCTTACATATATAGAAAGCGCTCCATTTACATATATGCAATGAAACATGATA 247
 DB 139 TCACTCTTACATATATAGAAAGCGCTCCATTTACATATATGCAATGAAACATGATA 198
 QY 248 CAGATTAAGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 307
 DB 199 CAGATTAAGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 258
 QY 308 CTGTAACTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 367
 DB 259 CTGTAACTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 318
 QY 368 GAAATTAACCAACCAATTAACCAACCAATTAACCAACCAATTAACCAATTAACCA 427
 DB 319 GAAATTAACCAACCAATTAACCAACCAATTAACCAACCAATTAACCAATTAACCA 378
 QY 428 CTCTTAAGTAAAGGAG 487
 DB 379 CTCTTAAGTAAAGGAG 438
 QY 488 AGGATTTCTTTGTCGTCATATGTTCCAAAGCGGTAACCTTGCAGCAGTAATACA 547
 DB 439 AGGATTTCTTTGTCGTCATATGTTCCAAAGCGGTAACCTTGCAGCAGTAATACA 498
 QY 548 CTGATGCTGTAACCGTAACCGTAATCTAACAAGATTCATGATTAGAGCC 597
 DB 499 CTGATGCTGTAACCGTAACCGTAATCTAACAAGATTCATGATTAGAGATCC 548

RESULT 9
 ID V41596 standard; DNA; 544 BP.
 AC V41596;
 DT 26-OCT-1998 (first entry)
 DE Nucleotide sequence of F1 antigen nyf1sec510.

KW F1 antigen; plasmid; vaccine; plague; ds.
 OS Yersinia pestis. Location/Qualifiers
 FH Key 17..532
 FT CDS
 FT /tag= a
 FT /product= "F1 antigen"
 FN WO9824912-A2.
 PD 11-JUN-1998.
 PF 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HEK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR WPI:96-33331/29.
 DR P-PSDB: W59783.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 8; Pages 53-54; 75pp; English.
 CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 CC Sequence 544 BP; 166 A; 118 C; 112 G; 148 T;
 SQ

Query Match 33.0%; Score 516.8; DB 1; Length 544;
 Best Local Similarity 99.6%; Pred. No. 1.8e-128;
 Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAAGTCCGTTATCGCATTCATTTTGGAACTATTCGACCTGCTA 127
 DB 15 ATATGAAAAAATCAAGTCCGTTATCGCATTCATTTTGGAACTATTCGACCTGCTA 74
 QY 128 ATGGCGCAGATTAACTGCAAGCAGCAGTCAAGGCACTCTGTGTAACGAGCCGCA 187
 DB 75 ATGGCGCAGATTAACTGCAAGCAGCAGTCAAGGCACTCTGTGTAACGAGCCGCA 134
 QY 188 TCACCTTACATATAGAGAGCGCTCCATTAATCAATTTGCAATGGAACATCGATA 247
 DB 135 TCACCTTACATATAGAGAGCGCTCCATTAATCAATTTGCAATGGAACATCGATA 194
 QY 248 CAGATTACTTGTGGTACGCTTACTCTGGGGGCTATTAACAGAACCACTAGACAT 307
 DB 195 CAGATTACTTGTGGTACGCTTACTCTGGGGGCTATTAACAGAACCACTAGACAT 254
 QY 308 CTGTAACTTACAGATGCGGGGATGCCATGACTTAATCAATTTACTCTCAGAGATG 367
 DB 255 CTGTAACTTACAGATGCGGGGATGCCATGACTTAATCAATTTACTCTCAGAGATG 314
 QY 368 GAAATTAACCAACATTAAGTATGGAAGGATCTAGAGATTTGATATCT 427
 DB 315 GAAATTAACCAACATTAAGTATGGAAGGATCTAGAGATTTGATATCT 374
 QY 428 CTCCTAAGTAAAGGTGGAACCTTGTGGGGATGAGTCTCTTGGCTAGGGGAGCC 487
 DB 375 CTCCTAAGTAAAGGTGGAACCTTGTGGGGATGAGTCTCTTGGCTAGGGGAGCC 434
 QY 488 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACTTGAGAGAGTAATACA 547
 DB 435 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACTTGAGAGAGTAATACA 494
 QY 548 CTGATGCTGTAAACCGTAAACGATATCAACCAAGATTCAT 587
 DB 495 CTGATGCTGTAAACCGTAAACGATATCAACCAAGATTCAT 534

RESULT 10
 T38248 standard; DNA; 547 BP.
 ID T38248;
 AC T38248;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1 antigen catI gene (including signal sequence).
 KW Plague; vaccine; genetic immunisation; F1 antigen; catI;

KW V antigen; ds. strain GB.
 OS Yersinia pestis. Location/Qualifiers
 FH Key signal_peptide 29..91
 FT mat_peptide 92..538
 FT /tag= a
 FT /tag= b
 FN WO9628551-A1.
 PD 19-SEP-1996.
 PF 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Titball RM, Williamson ED;
 DR WPI:96-433824/43.
 DR P-PSDB: W01043.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Claim 41; Page 61-62; 98pp; English.
 CC A DNA sequence (T38248) comprises the catI gene, including the
 CC signal sequence, coding for the F1 antigen (W01043) of Yersinia
 CC pestis. It was obtd. by PCR amplification (see also T38257-58)
 CC of Y. pestis DNA. The PCR product was cloned into plasmid pBKCMV
 CC and the resulting plasmid (pF1AB) was used to transform E. coli Nova
 CC Blue. Purified plasmid, when administered by i.m. injection,
 CC induced an immunoglobulin response to F1 in BALB/c mice. Live
 CC vaccines comprising gut colonising organisms transformed with the
 CC catI gene (see also T38244) can be used to protect a host animal
 CC against plague.
 CC Sequence 547 BP; 165 A; 120 C; 115 G; 147 T;
 SQ

Query Match 33.0%; Score 516.8; DB 1; Length 547;
 Best Local Similarity 99.6%; Pred. No. 1.8e-128;
 Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 68 ATATGAAAAAATCAAGTCCGTTATCGCATTCATTTTGGAACTATTCGACCTGCTA 127
 DB 27 ATATGAAAAAATCAAGTCCGTTATCGCATTCATTTTGGAACTATTCGACCTGCTA 86
 QY 128 ATGGCGCAGATTAACTGCAAGCAGCAGTCAAGGCACTCTGTGTAACGAGCCGCA 187
 DB 87 ATGGCGCAGATTAACTGCAAGCAGCAGTCAAGGCACTCTGTGTAACGAGCCGCA 146
 QY 188 TCACCTTACATATAGAGAGCGCTCCATTAATCAATTTGCAATGGAACATCGATA 247
 DB 147 TCACCTTACATATAGAGAGCGCTCCATTAATCAATTTGCAATGGAACATCGATA 206
 QY 248 CAGATTACTTGTGGTACGCTTACTCTGGGGGCTATTAACAGAACCACTAGACAT 307
 DB 207 CAGATTACTTGTGGTACGCTTACTCTGGGGGCTATTAACAGAACCACTAGACAT 266
 QY 308 CTGTAACTTACAGATGCGGGGATGCCATGACTTAATCAATTTACTCTCAGAGATG 367
 DB 267 CTGTAACTTACAGATGCGGGGATGCCATGACTTAATCAATTTACTCTCAGAGATG 326
 QY 368 GAAATTAACCAACATTAAGTATGGAAGGATCTAGAGATTTGATATCT 427
 DB 327 GAAATTAACCAACATTAAGTATGGAAGGATCTAGAGATTTGATATCT 386
 QY 428 CTCCTAAGTAAAGGTGGAACCTTGTGGGGATGAGTCTCTTGGCTAGGGGAGCC 487
 DB 387 CTCCTAAGTAAAGGTGGAACCTTGTGGGGATGAGTCTCTTGGCTAGGGGAGCC 446
 QY 488 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACTTGAGAGAGTAATACA 547
 DB 447 AGGATTTCTTTGTGCTCAATTTGTTCCAAAGCGGTAACTTGAGAGAGTAATACA 506
 QY 548 CTGATGCTGTAAACCGTAAACGATATCAACCAAGATTCAT 587
 DB 507 CTGATGCTGTAAACCGTAAACGATATCAACCAAGATTCAT 546

RESULT 11
ID V41595 standard; DNA: 510 BP.
AC V41595;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen nYP1(b)sec544.
KW F1 antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
PN W09824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
PT Protection of animals against plague - using nucleic acid encoding
PS antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Page 53; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SO Sequence 510 BP; 153 A; 114 C; 106 G; 137 T;

Query Match 32.6%; Score 510; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.1e-126;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 70 ATGAAAAAATCAGTCCGTTATGCCATTCGATATTTGGACATTTGCAACTGCTAAT 129
D 1 ATGAAAAAATCAGTCCGTTATGCCATTCGATATTTGGACATTTGCAACTGCTAAT 60
Y 130 GCGGAGATTTACTGCAAGCACCACTGCAAGCACTCTTTGTAACCAAGCCGCATC 189
D 61 GCGGAGATTTACTGCAAGCACCACTGCAAGCACTCTTTGTAACCAAGCCGCATC 120
Y 190 ACTCTTACATATAGGAAGGCGCTCCAAATTAATATGAGCAATGGAACATGATACA 249
D 121 ACTCTTACATATAGGAAGGCGCTCCAAATTAATATGAGCAATGGAACATGATACA 180
Y 250 GAATTAATCTGTTGACGCTTACTCTTGGCGGCTTAAACAGCAACCACTGCAATCT 309
D 181 GAATTAATCTGTTGACGCTTACTCTTGGCGGCTTAAACAGCAACCACTGCAATCT 240
Y 310 GTTAACCTTACAGATGCGGCGGATGCCATGTACTTACATTTACTTCTCAGATGGA 369
D 241 GTTAACCTTACAGATGCGGCGGATGCCATGTACTTACATTTACTTCTCAGATGGA 300
Y 370 AATAACCAACCAATTCATCAAAAGTGTGCAAGAGATTTAGAGATTTGATATCTCT 429
D 301 AATAACCAACCAATTCATCAAAAGTGTGCAAGAGATTTAGAGATTTGATATCTCT 360
Y 430 CCTAAGGTAAGCGGTGAGAACTTGTGGGGATGACGCTGTTGGCTACGGGCAAGCAG 489
D 361 CCTAAGGTAAGCGGTGAGAACTTGTGGGGATGACGCTGTTGGCTACGGGCAAGCAG 420
Y 490 GATTTCTTGTGCTCATTTGTTCCAAAGGCGGTAAGTGCAGAGGTAATATACACT 549
D 421 GATTTCTTGTGCTCATTTGTTCCAAAGGCGGTAAGTGCAGAGGTAATATACACT 480
Y 550 GATGCTGTAACCGTATACCGTATACCA 579
D 481 GATGCTGTAACCGTATACCGTATACCA 510

RESULT 12
ID V41600 standard; DNA: 474 BP.
AC V41600;
DT 26-OCT-1998 (first entry)

DE Nucleotide sequence of F1 antigen nYP1mat474.
KW F1 antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 7..459
FT /tag= a
FT /product= "F1 antigen"
PN W09824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR P-PSDB: W59787.
PT Protection of animals against plague - using nucleic acid encoding
PS antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 8; Pages 59-60; 75pp; English.
CC This is the nucleotide sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SO Sequence 474 BP; 143 A; 106 C; 99 G; 126 T;

Query Match 28.9%; Score 452.8; DB 1; Length 474;
Best Local Similarity 99.6%; Pred. No. 1.9e-111;
Matches 454; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Y 132 GGCAGATTTAATCGAACCACCACTGCAAGCGCACTCTTGTGACCAAGCCGATCAC 191
D 9 GGCAGATTTAATCGAACCACCACTGCAAGCGCACTCTTGTGACCAAGCCGATCAC 68
Y 192 TCTTACATATAGGAAGGCGCTCCAAATTAATATGAGCAATGGAACATGATACAGA 251
D 69 TCTTACATATAGGAAGGCGCTCCAAATTAATATGAGCAATGGAACATGATACAGA 128
Y 252 ATTACTTGTGTAAGCTTACTCTTGGCGGCTATTAACAGGAACCACTAGCACATCTGT 311
D 129 ATTACTTGTGTAAGCTTACTCTTGGCGGCTATTAACAGGAACCACTAGCACATCTGT 188
Y 312 TAACCTTACAGATGCGGCGGATGCCATGTACTTAAATTTACTTCTCAGATGGA 371
D 189 TAACCTTACAGATGCGGCGGATGCCATGTACTTAAATTTACTTCTCAGATGGA 248
Y 372 TAACCAACCAATTCATCAAAAGTGTGCAAGAGATTTGATATCTCTCC 431
D 249 TAACCAACCAATTCATCAAAAGTGTGCAAGAGATTTGATATCTCTCC 308
Y 432 TAAGTAACCGGTGAGAACTTGTGGGGATGACGCTGTTGGCTACGGGCAAGCAGA 491
D 309 TAAGTAACCGGTGAGAACTTGTGGGGATGACGCTGTTGGCTACGGGCAAGCAGA 368
Y 492 TTTCTTGTGCTCATTTGTTCCAAAGGCGGTAAGTGCAGAGGTAATATACACTGA 551
D 369 TTTCTTGTGCTCATTTGTTCCAAAGGCGGTAAGTGCAGAGGTAATATACACTGA 428
Y 552 TGCTGTAACCGTATACCGTATACCAAGATTCAT 587
D 429 TGCTGTAACCGTATACCGTATACCAAGATTCAT 464

RESULT 13
ID 092817
AC 092817 standard; DNA: 541 BP.
DT 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen in plasmid pFGAL2a.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
OS Yersinia pestis. Location/Qualifiers

FT cds 2..454
 FT /*tag- a
 FT misc-feature 1..6
 FT /*tag- b
 FT /note- "lacZ promoter fusion site"
 FT misc-feature 536..541
 FT /*tag- C
 FT /note- "vector pFAL2a bases"
 PN W09518231-A1.
 PD 06-JUL-1995.
 PE 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
 DR WPI; 95-246396/32.
 PT P-PSDB; R76526.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response,
 PT against Yersinia pestis, at mucosal surfaces.
 PS Claim 7; Page 15-16; 27pp; English.
 CC The sequence represents the plasmid pFAL2a construct showing the
 CC fusion of the first few bases of beta-galactosidase in the vector
 CC with the Y. pestis catf (F1) antigen minus its signal sequence and
 CC having a 5' tail including a SacI restriction site, and up to the
 CC catf AAC-3' end with some vector bases. The DNA construct can be
 CC used to transform human or animal gut colonizing microorganisms,
 CC specifically attenuated Salmonella typhimurium or Salmonella typh.
 CC The transformed microorganisms can be used as live/attenuated
 CC vaccines which induce immune responses at mucosal surfaces. The
 CC are parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 541 BP; 163 A; 120 C; 111 G; 147 T;

Query Match

Best Local Similarity 28.9%; Score 452.2; DB 1; Length 541;
 Matches 453; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 131 CGGCAGATTAACTGCAAGCAGCACTGCAAGGCACTCTGTTGAACGACCGCATCA 190
 DB 6 CGGAGATTAACTGCAAGCAGCACTGCAAGGCACTCTGTTGAACGACCGCATCA 65
 QY 191 CTCTACATATAAGGAAGCGCTCCCAATTACATTTATGGAATGGAACATGCATACAG 250
 DB 66 CTATTACATATAAGGAAGCGCTCCCAATTACATTTATGGAATGGAACATGCATACAG 125
 QY 251 AATTACTGTTGGTACGCTTACTCTTGGCGGCTATTAAGCAACGACCTGCACATCTG 310
 DB 126 AATTACTGTTGGTACGCTTACTCTTGGCGGCTATTAAGCAACGACCTGCACATCTG 185
 QY 311 TTAACCTTACAGATGCGCGGGGTATCCCATGTACTTAACTTACTCTCAGATGGA 370
 DB 186 TTAACCTTACAGATGCGCGGGGTATCCCATGTACTTAACTTACTCTCAGATGGA 245
 QY 371 ATAAACCAACATTCACATAAAGTATGGAAGATTCAGAGATTTGATATCTCTC 430
 DB 246 ATAAACCAACATTCACATAAAGTATGGAAGATTCAGAGATTTGATATCTCTC 305
 QY 431 CTAGGTAAGAGGTGAGAACTTGTGGGGGATGACGTCGTTGGCTACGCGGACCGAG 490
 DB 306 CTAGGTAAGAGGTGAGAACTTGTGGGGGATGACGTCGTTGGCTACGCGGACCGAG 365
 QY 491 ATTCTTGTGCTCATATGTTCTCAAGGCGGTAACTTGCAGAGATTAATACACTG 550
 DB 366 ATTCTTGTGCTCATATGTTCTCAAGGCGGTAACTTGCAGAGATTAATACACTG 425
 QY 551 ATGCTTAACCGTAACCGTATCTAACAAGATTCAT 587
 DB 426 ATGCTTAACCGTAACCGTATCTAACAAGATTCAT 462

RESULT 14

T38244
 ID T38244 standard; DNA; 541 BP.
 AC T38244;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis catf (F1 antigen) gene.
 KW Plague; vaccine; genetic immunisation; F1 antigen; catf;
 KW V antigen; ds.
 OS Yersinia pestis strain GB.
 FH Key
 FT misc_recomb location/Qualifiers
 FT 1..6
 FT /*tag- a
 FT /note- "includes the first few bases of the
 FT beta-galactosidase sequence in vector
 FT pFAL2a"
 FT 2..457
 FT cds
 FT /*tag- b
 FT /product- mature F1 antigen
 FT misc_recomb 536..541
 FT /*tag- C
 PN W09628551-A1.
 PD 19-SEP-1996.
 PE 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;
 DR WPI; 96-433824/43.
 PT P-PSDB; W01042.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Claim 7; Page 43-45; 98pp; English.
 CC The catf gene (T38244) codes for the Yersinia pestis F1 antigen
 CC (W01042), which is capable of evoking protective immune responses
 CC in animals. The gene was amplified from Y. pestis DNA by PCR using
 CC primers (T38245-46) homologous to the 5' and 3' ends of the gene.
 CC The catf PCR product was ligated into pUC18 and transformed into E.
 CC coli JM109 to produce vector pFAL2a. The catf gene can be used to
 CC express recombinant F1 antigen for use in vaccines against plague.
 CC Expression in gut-colonising organisms and attenuated Salmonella
 CC typh allows prodn. of live vaccines. The gene can itself be used
 CC in genetic vaccines. F1/V antigen fusions were also created (see
 CC also T38249 and T38256).
 SQ Sequence 541 BP; 163 A; 120 C; 111 G; 147 T;

Query Match

Best Local Similarity 28.9%; Score 452.2; DB 1; Length 541;
 Matches 453; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 131 CGGCAGATTAACTGCAAGCAGCACTGCAAGGCACTCTGTTGAACGACCGCATCA 190
 DB 6 CGGAGATTAACTGCAAGCAGCACTGCAAGGCACTCTGTTGAACGACCGCATCA 65
 QY 191 CTCTACATATAAGGAAGCGCTCCCAATTACATTTATGGAATGGAACATGCATACAG 250
 DB 66 CTATTACATATAAGGAAGCGCTCCCAATTACATTTATGGAATGGAACATGCATACAG 125
 QY 251 AATTACTGTTGGTACGCTTACTCTTGGCGGCTATTAAGCAACGACCTGCACATCTG 310
 DB 126 AATTACTGTTGGTACGCTTACTCTTGGCGGCTATTAAGCAACGACCTGCACATCTG 185
 QY 311 TTAACCTTACAGATGCGCGGGGTATCCCATGTACTTAACTTACTCTCAGATGGA 370
 DB 186 TTAACCTTACAGATGCGCGGGGTATCCCATGTACTTAACTTACTCTCAGATGGA 245
 QY 371 ATAAACCAACATTCACATAAAGTATGGAAGATTCAGAGATTTGATATCTCTC 430
 DB 246 ATAAACCAACATTCACATAAAGTATGGAAGATTCAGAGATTTGATATCTCTC 305
 QY 431 CTAGGTAAGAGGTGAGAACTTGTGGGGGATGACGTCGTTGGCTACGCGGACCGAG 490
 DB 306 CTAGGTAAGAGGTGAGAACTTGTGGGGGATGACGTCGTTGGCTACGCGGACCGAG 365

OY 491 ATTCTTGTTCGCTCAATTGGTTCACAAAGCGGTAACTTCAGCAGTAATACACTG 550
 DB 366 ATTCTTGTTCGCTCAATTGGTTCACAAAGCGGTAACTTCAGCAGTAATACACTG 425
 OY 551 ATGCTGTAAACCGTAACCGTATCTTAACCAAGAATTCAT 587
 DB 426 ATGCTGTAAACCGTAACCGTATCTTAACCAAGAATTCAT 462

RESULT 15

O92818 15
 ID Q92818 standard; DNA: 542 BP.

AC Q92818:
 DE 17-DEC-1995 (first entry)
 DE Yersinia pestis cafi (pt) antigen in plasmid pF313a.
 KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhl;
 OS bubonic plague; pneumonic plague; ds.
 OS Yersinia pestis.

Location/Qualifiers
 FH Key 3..455
 FT cds /tag- a

FT misc-feature 1..6

FT /tag- b /note- "E. coli LTB signal peptide bases"
 FT 536..541

FT misc-feature /tag- c

FT /note- "vector pF313a bases"

PN W09518231-A1.

PD 06-JUL-1995.

PE 23-DEC-1994; G02818.

PR 24-DEC-1993; GB-026425.

PA (MINA) UK SEC FOR DEFENCE.

PI Howells A. leary SEC, Oyston PCR, Tiltball RW, Williamson ED;

DR WPI: 95-246396/32.

DR P-PSDB; R/6527.

PT DNA constructs capable of transforming microorganisms - which can be

PT used as live or attenuated vaccines which induce an immune response,

PT against Yersinia pestis, at mucosal surfaces.

PS Disclosure: Page 17-18; 27pp; English.

CC The sequence represents the plasmid pF313a construct showing the

CC fusion of the first few bases of the E. coli LTB signal sequence

CC with the Y. pestis cafi (pt) antigen minus its signal sequence and

CC having a 5' tail including a SacI restriction site, and up to the

CC cafi AAC-3' end with some vector bases. The DNA construct can be

CC used to transform human or animal gut colonizing microorganisms,

CC specifically attenuated Salmonella typhimurium or Salmonella typhl.

CC The transformed microorganisms can be used as live/attenuated

CC vaccines which induce immune responses at mucosal surfaces. The

CC vaccines provide protection against infection with Y. pestis, and

CC are parenterally and orally active vaccines offering protection

CC against bubonic and pneumonic plague.

Sequence 542 BP; 163 A; 122 C; 110 G; 147 T;

Query Match 28.8%; Score 451; DB 1; Length 542;
 Best Local Similarity 98.9%; Pred. No. 5.9e-11;

Matches 454; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 129 TCGGCGAGATTAACTGCAAGCAGCAGTCAAGGCAACTTGTGTAACGACCCGCAT 188
 DB 5 TCGGCGAGATTAACTGCAAGCAGCAGTCAAGGCAACTTGTGTAACGACCCGCAT 64

OY 189 CACTTTCATATAAGAGAGGCGCTCCATTAATATATGACAAATGGAACATGCATAC 248
 DB 65 CACTTTCATATAAGAGAGGCGCTCCATTAATATATGACAAATGGAACATGCATAC 124

OY 249 AGAATTACTTGTGTAAGCTTACTCTGGGGCTATAAAGAGAAACCACTAGACATC 308
 DB 125 AGAATTACTTGTGTAAGCTTACTCTGGGGCTATAAAGAGAAACCACTAGACATC 184

OY 309 TGTTAACCTTACAGATGCCGGGTATCCATGTACTTAACATTTACTTCTCAGGATG 368
 DB 125 TGTTAACCTTACAGATGCCGGGTATCCATGTACTTAACATTTACTTCTCAGGATG 368

DB 185 TGTTAACCTTACAGATGCCGGGTATCCATGTACTTAACATTTACTTCTCAGGATG 244
 OY 369 AATAACCAACCAATTCACCTACAAAGATATGGCAAGATTCCTAAGATTTGATATCTC 428
 DB 245 AATAACCAACCAATTCACCTACAAAGATATGGCAAGATTCCTAAGATTTGATATCTC 304
 OY 429 TCCTAAGGTAAACGGGTAGAAACCTTGTGGGGATGACGCTGTGGCTAGCGGACCA 488
 DB 305 TCCTAAGGTAAACGGGTAGAAACCTTGTGGGGATGACGCTGTGGCTAGCGGACCA 364
 OY 489 GGATTTCTTTGTTCGCTCAATTGGTTCACAAAGCGGTAACTTCAGCAGGTAATACAC 548
 DB 365 GGATTTCTTTGTTCGCTCAATTGGTTCACAAAGCGGTAACTTCAGCAGGTAATACAC 424
 OY 549 TGATGCTGTAAACCGGTAAACCGGTATCTTAACCAAGAATTCAT 587
 DB 425 TGATGCTGTAAACCGGTAAACCGGTATCTTAACCAAGAATTCAT 463

Search completed: August 22, 2000, 15:36:39
 Job time: 2227 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2000, 14:58:41 ; Search time 46.08 Seconds
(without alignments)
4674.749 Million cell updates/sec

Title: US-08-699-716a-1

Sequence: 1566
1 ATGGCCATCATCATCATCA.....ATGACACGCTGTGAATGA 1566

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 243080 seqs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0
Maximum DB seq length: 100000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/lna/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/lna/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6_COMB.seq:*
6: /cgn2_6/ptodata/1/lna/PCBUS_COMB.seq:*
7: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1473.6	94.1	1530	4 US-08-913-477-22	Sequence 22, Appl
2	1409	90.0	1462	4 US-08-913-477-16	Sequence 16, Appl
3	976.8	62.4	1014	4 US-08-913-477-1	Sequence 1, Appl
4	974.8	62.2	1014	4 US-08-913-477-3	Sequence 3, Appl
5	513.6	32.8	547	4 US-08-913-477-20	Sequence 20, Appl
6	452.2	28.9	541	4 US-08-913-477-10	Sequence 10, Appl
7	68.4	4.4	375	5 US-08-870-370-7	Sequence 7, Appl
8	68.4	4.4	375	5 US-08-870-370-8	Sequence 8, Appl
9	68.4	4.4	375	5 US-08-870-370-9	Sequence 9, Appl
10	53.4	3.4	1402	2 US-08-480-604A-25	Sequence 25, Appl
11	53.4	3.4	1402	2 US-08-405-966A-25	Sequence 25, Appl
12	48	3.1	7218	1 US-08-232-463-14	Sequence 14, Appl
13	39	2.5	4435	1 US-08-308-872B-5	Sequence 5, Appl
14	38.6	2.5	1124	5 US-08-945-056-1	Sequence 11, Appl
15	37.4	2.4	4365	6 PCT-US93-08525-31	Sequence 31, Appl
16	37.4	2.4	4365	6 PCT-US93-04384-11	Sequence 11, Appl
17	35	2.2	47	4 US-08-600-783-15	Sequence 15, Appl
18	34.8	2.2	1446	4 US-08-560-007B-4	Sequence 4, Appl
19	34.8	2.2	1508	5 US-08-714-918-31	Sequence 31, Appl
20	34.8	2.2	2402	5 US-08-776-265-4	Sequence 4, Appl
21	34.8	2.2	2415	2 US-08-785-428-1	Sequence 1, Appl
22	34.8	2.2	2415	4 US-08-996-787-1	Sequence 1, Appl
23	34.6	2.2	1100	3 US-08-179-537-20	Sequence 20, Appl
24	34	2.2	867	4 US-09-036-582-36	Sequence 36, Appl
25	33.8	2.2	1947	5 US-08-604-991-1	Sequence 1, Appl
26	33.8	2.2	1947	5 US-09-363-639-1	Sequence 1, Appl

27	33.6	2.1	2902	5 US-08-714-918-95	Sequence 95, Appl
28	33.2	2.1	4565	5 US-08-776-265-1	Sequence 1, Appl
29	33.2	2.1	9636	2 US-08-323-170B-1	Sequence 1, Appl
30	33.2	2.1	10564	1 US-08-206-176-5	Sequence 5, Appl
31	33	2.1	1868	2 US-08-229-145-13	Sequence 13, Appl
32	33	2.1	2993	4 US-08-415-593-42	Sequence 42, Appl
33	32.8	2.1	8920	3 US-08-446-855A-1	Sequence 1, Appl
34	32.8	2.1	891	1 US-07-941-523-19	Sequence 19, Appl
35	32.8	2.1	1959	2 US-08-137-175A-1	Sequence 1, Appl
36	32.8	2.1	2369	3 US-08-525-742-1	Sequence 1, Appl
37	32.6	2.1	2653	3 US-08-589-711-1	Sequence 1, Appl
38	32.6	2.1	2794	1 US-07-960-932-1	Sequence 1, Appl
39	32.6	2.1	2794	1 US-07-908-253-1	Sequence 1, Appl
40	32.6	2.1	2794	2 US-08-455-970A-1	Sequence 1, Appl
41	32.6	2.1	2794	2 US-08-187-156-5	Sequence 5, Appl
42	32.6	2.1	2794	3 US-08-694-865-5	Sequence 5, Appl
43	32.6	2.1	2794	4 US-08-878-748-5	Sequence 5, Appl
44	32.6	2.1	2794	4 US-08-535-837-1	Sequence 1, Appl
45	32.6	2.1	2794	5 US-09-124-491-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-913-477-22
Sequence 22, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Titball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
CITY: Arlingtonton
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
TELECOMMUNICATION INFORMATION:
REFERENCE/DOCKET NUMBER: 124-599
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4000
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:

LENGTH: 1530 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Yersinia pestis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 13..1515
 US-08-913-477-22

Query Match 94.1% Score 1473.6; DB 4; Length 1530;
 Best Local Similarity 99.1% Pred. No. 0;
 Matches 1495; Conservative 0; Mismatches 4; Indels 9; Gaps 1;

QY 68 AATGTAAGAAAAATCAATGCTCCGTTATGCGCATTTGCAATTTATGGAAGTATGCAACTGCTA 127
 DB 11 AATGTAAGAAAAATCAATGCTCCGTTATGCGCATTTGCAATTTATGGAAGTATGCAACTGCTA 70
 QY 128 ATGGGCAATTTACTGCAAGCAGCAGTGCAGAGCGCACTCTGTTGAACCGCCGCA 187
 DB 71 ATGGGCAATTTACTGCAAGCAGCAGTGCAGAGCGCACTCTGTTGAACCGCCGCA 130
 QY 188 TCACTTCAATTAAGAAAGGCGCTCAATTCATTTATGCAATTTATGCAAGTATGCAACTGCTA 247
 DB 131 TCACTTCAATTAAGAAAGGCGCTCAATTCATTTATGCAATTTATGCAAGTATGCAACTGCTA 190
 QY 248 CAGAATTTACTGTTGTTAGCGTTACTCTTGGCGGTATTAAGAGAACAGACACATAGACAT 307
 DB 191 CAGAATTTACTGTTGTTAGCGTTACTCTTGGCGGTATTAAGAGAACAGACACATAGACAT 250
 QY 308 CTGTAACTTTACAGATGCGCGGGGGTGATGCCATGTTACTTAACTTACTCTTCAGAGATG 367
 DB 251 CTGTAACTTTACAGATGCGCGGGGGTGATGCCATGTTACTTAACTTACTCTTCAGAGATG 310
 QY 368 GAAATTAACCAACCAATTCATACAAAGATGTTGGCAAGATTTAGAGATTTTGAATCT 427
 DB 311 GAAATTAACCAACCAATTCATACAAAGATGTTGGCAAGATTTTGAATTTTGAATCT 370
 QY 428 CTCTTAAGTTAAAGCGTGAAGACCTTGGGGGATGACGTCGTTGGCTACGGGAGCC 487
 DB 371 CTCTTAAGTTAAAGCGTGAAGACCTTGGGGGATGACGTCGTTGGCTACGGGAGCC 430
 QY 488 AGGATTTCTTTGTTGCTCAATTTGTTCCAAAGCGGTAACTTGCAGCAGGTAAATACA 547
 DB 431 AGGATTTCTTTGTTGCTCAATTTGTTCCAAAGCGGTAACTTGCAGCAGGTAAATACA 490
 QY 548 CTGATGCTTAACCGTACCGTATCTACCAAGATTTCAAT-----GATTAGAGCGT 598
 DB 491 CTGATGCTTAACCGTACCGTATCTACCAAGATTTCAATCGAAGTTCGATTAGAGCGT 550
 QY 599 ACGAACAACCAACCAACATTTTATTAGAGATCTAGAAAAAGTTAGGTGGAACAACCTTA 658
 DB 551 ACGAACAACCAACCAACATTTTATTAGAGATCTAGAAAAAGTTAGGTGGAACAACCTTA 610
 QY 659 CTGGTCATGTTCTTCAAGTTTGAAGATTTGTTCAAGTTAGTCAAGTAAAGTAAATAG 718
 DB 611 CTGGTCATGTTCTTCAAGTTTGAAGATTTGTTCAAGTTAGTCAAGTAAAGTAAATAG 670
 QY 719 AATTTTCAATTAATTAATGATCCAGAAAGATCGAGAGTTTGGCAATAGAGTAATA 778
 DB 671 AATTTTCAATTAATTAATGATCCAGAAAGATCGAGAGTTTGGCAATAGAGTAATA 730
 QY 779 CTGATGATATCGAATTTGTCAGAAAAATCTTACTTATTTTCTACCCGAGATACCATTC 838
 DB 731 CTGATGATATCGAATTTGTCAGAAAAATCTTACTTATTTTCTACCCGAGATACCATTC 790
 QY 839 TTAAGGCGGTCTTTGACAAACCACTGCAAAATGTCATCAAGCAGTAAAGAGTTCC 898

DB 791 TTAAGGCGGTCTTTGACAAACCAACTGCAAAATGCGATCAAGCGAGTAAAGAGTTCC 850
 QY 899 TTGAATCATCGCCGAATACACATGGGAATTTGCGGGCGTTCAAGGAGTAAATGCAATTC 958
 DB 851 TTGAATCATCGCCGAATACACATGGGAATTTGCGGGCGTTCAAGGAGTAAATGCAATTC 910
 QY 959 CTTTAACCGCGCATGATGATGATATTTTGAAGTATTTGATTTGATTTCAATGAATC 1018
 DB 911 CTTTAACCGCGCATGATGATGATATTTTGAAGTATTTGATTTGATTTCAATGAATC 970
 QY 1019 ATCATGATATGCGGTGACAGAGTTGCGTGAAGATTTAGCTAGCTTACCGCCGAATTA 1078
 DB 971 ATCATGATATGCGGTGACAGAGTTGCGTGAAGATTTAGCTAGCTTACCGCCGAATTA 1030
 QY 1079 AGATTTATCAATTTATTCAGCCGAATTAATAGATCTGCTAGTATGSCACACATA 1138
 DB 1031 AGATTTATCAATTTATTCAGCCGAATTAATAGATCTGCTAGTATGSCACACATA 1090
 QY 1139 AATCCATGATTAATTCATTAATCTCATGATGATTAATTAATGATGATTAATGATGAAG 1198
 DB 1091 AATCCATGATTAATTCATTAATCTCATGATGATTAATTAATGATGATTAATGATGAAG 1150
 QY 1199 AGATTTTAAAGCCGAGAGTACAAATTTCTGAGAAAAATGCTCAACCCCATTC 1258
 DB 1151 AGATTTTAAAGCCGAGAGTACAAATTTCTGAGAAAAATGCTCAACCCCATTC 1210
 QY 1259 AGTGTATGAGAGAGAAAAAATAGTCTGATTAAGAGCTTCTTGGAAATGAGATA 1318
 DB 1211 AGTGTATGAGAGAGAAAAAATAGTCTGATTAAGAGCTTCTTGGAAATGAGATA 1270
 QY 1319 AAGAAGCGGGCGTTGGTAAATCTGAATACTCATCTTATATAAAGTAAATAG 1378
 DB 1271 AAGAAGCGGGCGTTGGTAAATCTGAATACTCATCTTATATAAAGTAAATAG 1330
 QY 1379 AATATCTACTTTGGCACCACTGCTCGGATTAAGTCCAGGCGGCTCAACGACTTGTTA 1438
 DB 1331 AATATCTACTTTGGCACCACTGCTCGGATTAAGTCCAGGCGGCTCAACGACTTGTTA 1390
 QY 1439 GCCAAAAACAACCTGAGCTGATATATACATCAAGTTTAACTCAGTATTTGAAGC 1498
 DB 1391 GCCAAAAACAACCTGAGCTGATATATACATCAAGTTTAACTCAGTATTTGAAGC 1450
 QY 1499 TGAACGTTTCAATGAAATATGATTCAGTATGATCAACGCTGATGATGACAGTCTG 1558
 DB 1451 TGAACGTTTCAATGAAATATGATTCAGTATGATCAACGCTGATGATGACAGTCTG 1510
 QY 1559 GTTAAATGA 1566
 DB 1511 GTTAAATGA 1518

RESULT 2
 US-08-913-477-16
 ; Sequence 16, Application US/08913477
 ; Patent No. 5985285
 ; GENERAL INFORMATION:
 ; APPLICANT: Tibball, Richard W.
 ; APPLICANT: Williamson, Ethel D.
 ; APPLICANT: Leary, Sophie E.C.
 ; APPLICANT: Oyston, Petra C.F.
 ; APPLICANT: Bennett, Alice M.
 ; TITLE OF INVENTION: VACCINES FOR PLAGUE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHYE P.C.
 ; STREET: 1100 No. 5985285th Glebe Rd. 8th floor
 ; CITY: Arlington
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22201-4741
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1462 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 8..1447
US-08-913-477-16

```

```

Query Match          90.0%; Score 1409; DB 4; Length 1462;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1431; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

```

```

QY 131 CGCGAGATTAACTGCAAGCACCACCTGCAAGGCAACTCTTGTGAACGACCCGCATCA 190
DB 6 CGCGAGATTAACTGCAAGCACCACCTGCAAGGCAACTCTTGTGAACGACCCGCATCA 65
QY 191 CTTTACATATATAGGAAGGCGCTCCCAATTACATTTATGACATATGAAACATCGATACG 250
DB 66 CTTTACATATATAGGAAGGCGCTCCCAATTACATTTATGACATATGAAACATCGATACG 125
QY 251 AATTACTTGTGTAGCTTACTCTTGGCGGCTATAAAGAGAAACACCTAGACATCTG 310
DB 126 AATTACTTGTGTAGCTTACTCTTGGCGGCTATAAAGAGAAACACCTAGACATCTG 185
QY 311 TTAATCTTACAGATGCGGCGGCTGATCCATGTACTTAACATTACTTCTCAGATGAA 370
DB 186 TTAATCTTACAGATGCGGCGGCTGATCCATGTACTTAACATTACTTCTCAGATGAA 245
QY 371 ATATACCAATTTCACTATAAAGATATGAGCAAGATTTCTAGATTTGATATCTTC 430
DB 246 ATATACCAATTTCACTATAAAGATATGAGCAAGATTTCTAGATTTGATATCTTC 305
QY 431 CTAAAGTAACGCTGAGAACTTGTGGGGATGACGCTGCTGTGGCTAGCGGACGACG 490
DB 306 CTAAAGTAACGCTGAGAACTTGTGGGGATGACGCTGCTGTGGCTAGCGGACGACG 365
QY 491 ATTCTTTTGTGCTCAATGTGTTCCAAAGCGGTAAACTTGCAGCAGGTAAATACACTG 550
DB 366 ATTCTTTTGTGCTCAATGTGTTCCAAAGCGGTAAACTTGCAGCAGGTAAATACACTG 425

```

```

QY 551 ATGCTTAACCGTAACCGGTATCTAACCAAGAAATTCAT-----GATTAGAGCCTAGC 601
DB 426 ATGCTTAACCGTAACCGGTATCTAACCAAGAAATTCAT-----GATTAGAGCCTAGC 485
QY 602 AACAAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGGGTGAACAACATCTAGC 661
DB 486 AACAAAACCCACACATTTTATGAGATCTAGAAAAAGTTAGGGTGAACAACATCTAGC 545
QY 662 GTATGCTTCTTCACTTTTGAAGAATTGCTCAGTTAGTCAAGATTAATAATATAGATA 721
DB 546 GTATGCTTCTTCACTTTTGAAGAATTGCTCAGTTAGTCAAGATTAATAATATAGATA 605
QY 722 TTTTCATTAATATGATCCCAAAAAAGATGAGAGTTTTCCTCAATAGATAATAGC 781
DB 606 TTTTCATTAATATGATCCCAAAAAAGATGAGAGTTTTCCTCAATAGATAATAGC 665
QY 782 ATGATATCGAATTTGCTCAAGAAATCCATTTATTTCTACCCGAGATACCATCTTA 841
DB 666 ATGATATCGAATTTGCTCAAGAAATCCATTTATTTCTACCCGAGATACCATCTTA 725
QY 842 AAGCGGCTATTATGACAAACCACTGCAAAATGCAATGCAAGCGAGTAAAGAAGTCTTG 901
DB 726 AAGCGGCTATTATGACAAACCACTGCAAAATGCAATGCAAGCGAGTAAAGAAGTCTTG 785
QY 902 AATCATCGCGCAATATACAAATGGAATGCGGCGCTCATGCGCAGTAATGCTCTCT 961
DB 786 AATCATCGCGCAATATACAAATGGAATGCGGCGCTCATGCGCAGTAATGCTCTCT 845
QY 962 TAACCGCGCATGTATGATGATGATTTTGAAGTGAATGTTGATTAATGAATCATC 1021
DB 846 TAACCGCGCATGTATGATGATGATTTTGAAGTGAATGTTGATTAATGAATCATC 905
QY 1022 ATGCTGATCCCGTACAGAAATGCTGGAAGAAATGCTGAGCTTACCGCGCAATTAAGA 1081
DB 906 ATGCTGATCCCGTACAGAAATGCTGGAAGAAATGCTGAGCTTACCGCGCAATTAAGA 965
QY 1082 TTTTATCATTTATTCAGCGGAAATTAATACATCTGCTAGTAGTGGCACCATTAATA 1141
DB 966 TTTTATCATTTATTCAGCGGAAATTAATACATCTGCTAGTAGTGGCACCATTAATA 1025
QY 1142 TCCATGATTAATTCATTAATCTCATGATTAATAATTTATATGTTATACAGATGAAGA 1201
DB 1026 TCCATGATTAATTCATTAATCTCATGATTAATAATTTATATGTTATACAGATGAAGA 1085
QY 1202 TTTTAAAGCCAGCGCAGATCAAAATTTCTGAGAAAAATGCTCAAAACCATTCAGG 1261
DB 1086 TTTTAAAGCCAGCGCAGATCAAAATTTCTGAGAAAAATGCTCAAAACCATTCAGG 1145
QY 1262 TGGATGGAGCGAGAAAAAATAGTCTGATTAAGGACTTCTTGGAGTGAAGATAAA 1321
DB 1146 TGGATGGAGCGAGAAAAAATAGTCTGATTAAGGACTTCTTGGAGTGAAGATAAA 1205
QY 1322 GAACGGGCGCTGGGTAATCTGAATAACTCATACTCTATATATAAGATTAATTAATGAAT 1381
DB 1206 GAACGGGCGCTGGGTAATCTGAATAACTCATACTCTATATATAAGATTAATTAATGAAT 1265
QY 1382 TATCTACTTTGCCACACACCTGCTGGATAGTCCAGGCGCTCAACGACTTGTTAGCC 1441
DB 1266 TATCTACTTTGCCACACACCTGCTGGATAGTCCAGGCGCTCAACGACTTGTTAGCC 1325
QY 1442 AAAAAAACAACCTAGCTGTGTGATTAATACACGTTTAAATACAGTATTAAGACACTGA 1501
DB 1326 AAAAAAACAACCTAGCTGTGTGATTAATACACGTTTAAATACAGTATTAAGACACTGA 1385
QY 1502 ACCGTTTCATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1561
DB 1386 ACCGTTTCATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1445
QY 1562 AATGA 1566
DB 1446 AATGA 1450

```

RESULT 3
US-08-913-477-1
Sequence 1, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibbail, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 No. 5985285th Glebe Rd. 8th floor
City: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1014 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..987
US-08-913-477-1

Query Match 62.48; Score 976.8; DB 4; Length 1014;
Best Local Similarity 99.88; Pred. No. 5.9e-276;
Matches 978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 587 TGATTAGAGCTTACGACAAACCCACAAATTTATGAGAGCTGAAAAAGTTAGGG 646
11 TCATTAGAGCTTACGACAAACCCACAAATTTATGAGAGCTGAAAAAGTTAGGG 70

QY 647 TGAACAACCTACTGTCATGTTCTTCACTTTAGAGAATTGGTTCAGTACGTCAG 706
DB 71 TGAACAACCTACTGTCATGTTCTTCACTTTAGAGAATTGGTTCAGTACGTCAG 130
QY 707 ATAAAAATATGATATTTCCATTAAATATGATCCGAGAAAAGATCGAGAGTTTCCCA 766
DB 131 ATAAAAATATGATATTTCCATTAAATATGATCCGAGAAAAGATCGAGAGTTTCCCA 190
QY 767 ATAGATTAATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 826
DB 191 ATAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250
QY 827 AGATACCACTTTTAAAGCGGCTATTATGACAACCACTGCAAAATGCGATCAAGCGAG 886
DB 251 AGATACCACTTTTAAAGCGGCTATTATGACAACCACTGCAAAATGCGATCAAGCGAG 310
QY 887 TAAAGAGTCTCTGATCATCGCCGAAATACAAATGGGAATTGCGGCGTTATGCGCAG 946
DB 311 TAAAGAGTCTCTGATCATCGCCGAAATACAAATGGGAATTGCGGCGTTATGCGCAG 370
QY 947 TAATGATTTCTCTTTAAACCGCGATCGTATGATGATGATGATGATGATGATGATGATG 1006
DB 371 TAATGATTTCTCTTTAAACCGCGATCGTATGATGATGATGATGATGATGATGATGATG 430
QY 1007 ATTCAATGAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
DB 431 ATTCAATGAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490
QY 1067 CCGCGCAATTAAGATTTATTCAGTTTATCAAGCCGAAATTAATGACATCTGCTAGTA 1126
DB 491 CCGCGCAATTAAGATTTATTCAGTTTATCAAGCCGAAATTAATGACATCTGCTAGTA 550
QY 1127 GTGCGACCATTAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
DB 551 GTGCGACCATTAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610
QY 1187 ATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTGAGAAAATGCTC 1246
DB 611 ATACAGATGAAGAGATTTTAAAGCCAGCGCAGAGTACAAATTTCTGAGAAAATGCTC 670
QY 1247 AAACCCACATTCAGTGTGAGTGGGAGCGCAGAAAAAATAGCTGATTAAGAGCTTTCTG 1306
DB 671 AAACCCACATTCAGTGTGAGTGGGAGCGCAGAAAAAATAGCTGATTAAGAGCTTTCTG 730
QY 1307 GAATGAGATTAATTAAGAACCGGCGGCTGGGTATCTGAATAAATCTACTACTTATATA 1366
DB 731 GAATGAGATTAATTAAGAACCGGCGGCTGGGTATCTGAATAAATCTACTACTTATATA 790
QY 1367 AAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1426
DB 791 AAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 850
QY 1427 ACGACTGTGTTAGCCAAAAAACAACCTAGCTGCTGATTAATTAATTAATTAATTAAT 1486
DB 851 ACGACTGTGTTAGCCAAAAAACAACCTAGCTGCTGATTAATTAATTAATTAATTAAT 910
QY 1487 CTATTGAAGCACTGACCGCTTCAATCAGAAATATGATGATGATGATGATGATGATGAT 1546
DB 911 CTATTGAAGCACTGACCGCTTCAATCAGAAATATGATGATGATGATGATGATGATGAT 970
QY 1547 ATGACAGCTGTGTAATGA 1566
DB 971 ATGACAGCTGTGTAATGA 990

RESULT 4
US-08-913-477-3
Sequence 3, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibbail, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.

APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Gleebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1014 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..987
US-08-913-477-3

Query Match 62.2%; Score 974.8; DB 4; Length 1014;
Best Local Similarity 99.8%; Pred. No. 2.3e-275;
Matches 976; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 589 ATTAGAGCTTACGACAAACCCACACATTTATTGAGAGCTAGAAAAGTTAGGGG 648
DB 13 ATTGAGCTTACGACAAACCCACACATTTATTGAGAGCTAGAAAAGTTAGGGG 72
QY 649 GACACACTTACGTCAGTCAGTTCTTCAGTTTAGAAGAAATTGGTTCAGTCAAGAT 708
DB 73 GACACACTTACGTCAGTCAGTTCTTCAGTTTAGAAGAAATTGGTTCAGTCAAGAT 132
QY 709 AAAATATGATATTTCCATTAAATATGATCCAGAAAAGATTGGAGGTTTGGCAAT 768
DB 133 AAAATATGATATTTCCATTAAATATGATCCAGAAAAGATTGGAGGTTTGGCAAT 192
QY 769 AGAGTAATTAATGATGATATGCAATGTCACAGAAAATCTAGCTTATTTCTACCGAG 828

DB 193 AGAGTAATTAATGATGATATGCAATGTCACAGAAAATCTAGCTTATTTCTACCGAG 252
QY 829 GATACACTTCTTAAAGGGGTATATATGACAAACCACTGCAAAATGGCATCAAGGAGTA 888
DB 253 GATGCAATCTTAAAGGGGTATATATGACAAACCACTGCAAAATGGCATCAAGGAGTA 312
QY 889 AAAGATTCCTTGAATCATGCGCGAATACACAAATGGGAATTCGGGCGTCTGAGCAGTA 948
DB 313 AAAGATTCCTTGAATCATGCGCGAATACACAAATGGGAATTCGGGCGTCTGAGCAGTA 372
QY 949 ATGCAATTCCTTAAACCGCGATGATGATGATGATGATGATGATGATGATGATGAT 1008
DB 373 ATGCAATTCCTTAAACCGCGATGATGATGATGATGATGATGATGATGATGATGAT 432
QY 1009 TCAATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
DB 433 TCAATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
QY 1069 GCCGAATTAAGATTTATTCAGTATTCAGTATTCAGTATTCAGTATTCAGTATTCAGT 1128
DB 493 GCCGAATTAAGATTTATTCAGTATTCAGTATTCAGTATTCAGTATTCAGTATTCAGT 552
QY 1129 GGCACATTAATTAATTCATGATTAATTCATGATTAATTCATGATTAATTCATGAT 1188
DB 553 GGCACATTAATTAATTCATGATTAATTCATGATTAATTCATGATTAATTCATGAT 612
QY 1189 ACAGATGAAGAGATTTTAAACCCAGCGAGATGACAAATTCGAGAAAATGCTGCA 1248
DB 613 ACAGATGAAGAGATTTTAAACCCAGCGAGATGACAAATTCGAGAAAATGCTGCA 672
QY 1249 ACCACATTCAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
DB 673 ACCACATTCAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
QY 1309 AGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1368
DB 733 AGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 792
QY 1369 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1428
DB 793 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 852
QY 1429 GACTGTTAGCCAAAACCAACCTGCTGATTAATTAATTAATTAATTAATTAATTAAT 1488
DB 853 GACTGTTAGCCAAAACCAACCTGCTGATTAATTAATTAATTAATTAATTAATTAAT 912
QY 1489 ATTGAAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTATTCAGTATTCAGT 1548
DB 913 ATTGAAGCACTGAACCGTTTCATTCAGAAATATGATTCAGTATTCAGTATTCAGT 972
QY 1549 GACACGTCGTGAATGA 1566
DB 973 GACACGTCGTGAATGA 990

RESULT 5
US-08-913-477-20
Sequence 20, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tibball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 No. 5985285th Gleebe Rd. 8th floor
CITY: Arlington
STATE: VA

COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis
FEATURE:
NAME/KEY: CDS
LOCATION: 29..538
US-08-913-477-20

Query Match 32.8%; Score 513.6; DB 4; Length 547;
Best Local Similarity 99.2%; Pred. No. 8.8e-141;
Matches 516; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 68 ATATGAAAAAATCAGTCCGTTATCGCATTCATATTTGGAACTATGCAACTGCTA 127
DB 27 ATATGAAAAAATCAGTCCGTTATCGCATTCATATTTGGAACTATGCAACTGCTA 86
QY 128 ATCGGCGAGATTAACTGCAAGCACCCTGCAACGGCACTCTTGTGGAACGACCGCA 187
DB 87 ATCGGCGAGATTAACTGCAAGCACCCTGCAACGGCACTCTTGTGGAACGACCGCA 146
QY 188 TCACCTCTTACATTAAGGAAGGGGCTTCATTAATTAAGGACAAATGCAATGATA 247
DB 147 TCACCTCTTACATTAAGGAAGGGGCTTCATTAATTAAGGACAAATGCAATGATA 206
QY 248 CAGAAATTAATTTGGTACGCTTACTCTTGGCGCTATAAAGCAAGCACTAGACAT 307
DB 207 CAGAAATTAATTTGGTACGCTTACTCTTGGCGCTATAAAGCAAGCACTAGACAT 266
QY 308 CTGTAACTTACAGATGCCGCGGGTGATCCATGTAATTAATTAATTTACTTCAGGATG 367
DB 267 CTGTAACTTACAGATGCCGCGGGTGATCCATGTAATTAATTAATTTACTTCAGGATG 326
QY 368 GAAATAACCAACCAATTCATCAAAAGTATTGGCAAGGATTCTAGAGATTGATATCT 427

DB 327 GAAATAACCAACCAATTCATCAAAAGTATTGGCAAGGATTCTAGAGATTGATATCT 386
QY 428 CTCCTAAGTAACGGGTGAGAACCTTGTGGGGATGACGCTGTGGCTAGCGGACGC 487
DB 387 CTCCTAAGTAACGGGTGAGAACCTTGTGGGGATGACGCTGTGGCTAGCGGACGC 446
QY 488 AGGATTTCTTTGTCCTCAATTTGTTCCAAAGCGCGTAACCTTGACGAGGTAATAACA 547
DB 447 AGGATTTCTTTGTCCTCAATTTGTTCCAAAGCGCGTAACCTTGACGAGGTAATAACA 506
QY 548 CTGATGCTGTAAACCGTAACCGTATCTAACCAGAAATTCAT 587
DB 507 CTGATGCTGTAAACCGTAACCGTATCTAACCAGAAATTCAT 546
RESULT 6
US-08-913-477-10
Sequence 10, Application US/08913477
Patent No. 5985285
GENERAL INFORMATION:
APPLICANT: Tilball, Richard W.
APPLICANT: Williamson, Ethel D.
APPLICANT: Leary, Sophie E.C.
APPLICANT: Oyston, Petra C.F.
APPLICANT: Bennett, Alice M.
TITLE OF INVENTION: VACCINES FOR PLAGUE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYTE P.C.
STREET: 1100 No. 5985285th Glade Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,477
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00571
FILING DATE: 13-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9505059.7
FILING DATE: 13-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9518946.0
FILING DATE: 15-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9524825.8
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4000
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 541 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Yersinia pestis

RESULT 13

US-08-308-872B-5
; Sequence 5, Application US/08308872B
; Patent No. 5651006

GENERAL INFORMATION:

APPLICANT: BROWN, Thomas David Kay
APPLICANT: HORSBURGH, Brian Collin
TITLE OF INVENTION: CANINE CORONAVIRUS SUBUNIT VACCINE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Akzo No. 5661006el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,872B

FILING DATE: 19-SEP-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,846

FILING DATE: 05-APR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,641

FILING DATE: 24-APR-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91.303.737.0

FILING DATE: 25-APR-1991

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.

REGISTRATION NUMBER: 34,409

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200

INFORMATION FOR SEQ. ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 4435 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Canine corona virus

STRAIN: CCV-V54

FEATURE:

NAME/KEY: CDS

LOCATION: 60..4418

OTHER INFORMATION: /label=CCV-c54_Spikegene

US-08-308-872B-5

Query Match 2.5%; Score 39; DB 1; Length 4435;

Best Local Similarity 48.8%; Pred. No. 0.12; Indels 0; Gaps 0;

Matches 105; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 622 ATTGAGATCTAGAAAAGTTAGGCTGGAACAATTACTGTCATGCTTTCAGTTTAA 681
DB 120 AATAAAGATGTGTACAAAGTAATGTACACAAATTCCTGCAATGAAATATCATTTAA 179
QY 682 GAAGAATGTTTACGTTAGTCAAGAAAGATAAATATAGATATTCATTAATATGATCCC 741
DB 180 GATTTCCTAATAGATTTTAAAGAGAGAGTGTAGTTGTTGGCTACTACCCC 239
QY 742 AGAAAGATTCGGAGGTTTGGCAATAGAGTATTTCTGATGATGATGCAATGCTCAAG 801
DB 240 ACAGAGGTGTGTACACATGTTCCAGACAGCAACACTACAGCTTACCTATTATTAGT 299

RESULT 14

US-08-945-056-1
; Sequence 1, Application US/08945056
; Patent No. 6077994

GENERAL INFORMATION:

APPLICANT: Coupland, George M.
APPLICANT: Putterill, Joanna J.
TITLE OF INVENTION: Genetic control of flowering
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSER: Nixon & Vanderhye PC
STREET: 8th Floor, 1100 No. 6077994th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,056

FILING DATE: 20-OCT-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02561

FILING DATE: 01-NOV-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9422083.7

FILING DATE: 02-NOV-1994

ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 620-17

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1124 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Arabidopsis thaliana

STRAIN: Landsberg erecta

POSITION IN GENOME:

CHROMOSOME/SEGMENT: chromosome 5

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1119

US-08-945-056-1

Query Match 2.5%; Score 38.6; DB 5; Length 1124;

Best Local Similarity 52.1%; Pred. No. 0.075; Indels 0; Gaps 0;

Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1 ATGGGCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
DB 664 AGGGGCGACAGTGCATTAACCAAGAAATTTTCAGTTCATTAATCAATATGAGCTCCTCA 723
QY 61 GACAACATATGAAAAAATCAAGTTCGTTATTCGCTATTCATTCATTCATTCATTCATTC 120
DB 724 GAGCACTACTACACAGCAACATGTTCCATTAACCAATACGCAATACATTCATTCATTC 783

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 22, 2000, 14:56:42 ; Search time 613.89 Seconds
(without alignments)
11249.232 Million cell updates/sec

Title: US-08-699-716a-1

Perfect score: 1566
Sequence: 1 ATGGGCCATCATCATCA.....ATGACAGCTCTGTAATGA 1566

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_est56:*
94: gb_est57:*
95: gb_est58:*
96: gb_est59:*
97: em_est38:*
98: em_est39:*
99: em_est40:*
100: em_est41:*
101: gb_est60:*
102: gb_est61:*
103: gb_est62:*
104: gb_est63:*
105: gb_est64:*
106: em_est60:*
107: em_est61:*
108: em_est62:*
109: em_est63:*
110: em_est64:*
111: em_est65:*
112: em_est66:*
113: gb_est67:*
114: gb_est68:*
115: em_est69:*
116: gb_est70:*

```

117:  gb_gss13:*
118:  gb_gss14:*
119:  gb_gss15:*
120:  gb_gss16:*
121:  gb_gss17:*
122:  gb_gss18:*
123:  gb_gss19:*
124:  em_gss13:*

```

cy
ch

ALIGNMENTS

ACCESSION fly), genomic survey sequence.
AL069652
VERSION AL069652.1 GI:4949997
KEYWORDS GSS.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information, please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPcI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
SOURCE location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_1lb="RPcI-98"
/clone="BACR30E19"
/note="end : TTTT"

BASE COUNT 300 a 161 c 209 g 167 t 264 others

ORIGIN

Query Match 3.2%; Score 50.6; DB 122; Length 1101;
Best Local Similarity 36.1%; Pred. No. 0.0063;
Matches 84; Conservative 49; Mismatches 100; Indels 0; Gaps 0;

OY 610 CCACAACATTTTATTGAGGATCTAGAAAAGTTAGGGTGGAACAATTACTGTGCATCGT 669
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 967 CAAAMWGAAGCAATTKAAYVAHAATAATTTTHMTTTTTTAAGVCMCHMHAATTTWTAAKT 908

OY 670 TCTTCAGTTTNGAAGAATGGTTCAGTACTACCAAGATAAATAATATGATATTTCCATT 729
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 907 TTTTITTTTTCACMTTGMWKTKTGTGTTTTTAAATMAAAAAAATAATDMAAAAWTT 848

OY 730 AAATATGATCCAGAAAAGATTTCGAGGTTTTTGGCATTAGAGTAATTAATCATGATGATATC 789
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 847 AAMWWMAAACCMHMBMTAAACTCGMTGTTTTCCKOAAAAGARAGAAAAACAAMGAWMW 788

OY 790 GAATTCGTCAGAAAATCCTTAGCTTATTTTCTACC CGAGGATACCATTTCTAA 842
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 787 CWTWMDAAACWCAMAAMAMAMAMATTTTAACMAAAGAMTTCATTGMAA 735

RESULT 2
CNSD0396 LOCUS 1101 bp DNA GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR0BK10 of RPcI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL069652
VERSION AL069652.1 GI:4949997
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

Accession	Definition	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal	Comment
CNS00396/C	CNS00396	1101 bp	DNA	GSS	03-JUN-1999					
	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPCT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.									
AL063921	AL063921	GI:4941778								
SS.	fruit fly									
Drosophila	Drosophila melanogaster									
Eutaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.										
(bases 1 to 1101)										
Genoscope.										
Direct Submission										
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage										
Bp 191 91006 EVRY cedex - FRANCE (E-mail : segr@genoscope.cns.fr										
- Web : www.genoscope.cns.fr)										
Determination of this BAC end and sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).										
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oosawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .										
FEATURES										
Source										
1..1101										
/organism="Drosophila melanogaster"										
/db_xref="taxon:7227"										
/clone_lib="RPCT-98"										
/clone="BACR08K10"										
/note="end : TET3"										
BASE COUNT	201 a	64 c	131 g	202 t	503 others					
ORIGIN										
Query Match	3.0%	Score 47;	DB 122;	Length 1101;						
Best Local Similarity	14.1%	Pred. Mismatches	0.057;							
Matches	59;	Conservative	198;	No. matches	162;	Indels	0;	Gaps	0;	
974	GTATCATGATGATGATTTTGAAGTAGTATGTTGATTCATGATCATGATCATCTCATGGGATGCC	1033								
946	KKATKMWDDWDDKDKKMGAKDRADDGAGDKDDGKGDADDGTDGTDKDDDKDX	887								
1034	GTAGCAAGTTGGCGTGAAGATTAGCTGACCTTACCCGCCGATTAAGATTATTCAGTA	1093								
886	DDMDRAKGTGADATWAAATDMMWGMADADMMWTDAAADMMWADRWDMAMKMDANA	827								
1094	TTCAAGCCGAATTAATTAAGCATCTGTCTAGTAGTGCCACCATTAATTCATGATAAAT	1153								
826	WGAFTADRRDMDWRACKRGGARRRRDRKADRDADRDADAATTTTTRTPTDWM	767								
1154	CCATTAATCTCATGGATTAATAATTTATGTTATGATATACATGATGAAGATTTTAAAGCA	1213								
766	KKKTDTWTWMAADRTWRDDDDDDRDAGTAGRKRRRTTTRRRKRRRTTWRDDADADDTAD	707								
1214	GCGCAGATGACAAATTCGAGAAATCCCTCAACCAACCATTCAGTGATGGAGCG	1273								
706	DRRRGGDDGADAGKGGKTKRRRRRDRATWDRTDAMWADAATTTTDTDDDKDRRR	647								
1274	AGAAAAAATAGTCTGCTGATAAAGGACTTCTTGGAAGTAGAATAAAGAACCGGGCGT	1333								
646	KCARRRRTTAAGAAWMTWTKAMDWAKMCKTFADRWDAADTTWTDARKADRWAKARA	587								

QY	1334	TGGGAAATCTGAAATCAATCTCTTTATTAATAAGAAATATATCTACTCTT	1392
Db	566	WRARDDRAARAARADRRWTTTKGTTATATTTAAARAAMAMAMATTTATTTTWT	528
RESULT	5		
LOCUS	CNS003F8		
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BAR08C03 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL064634		
VERSION	AL064634.1		
KEYWORDS	GSS.		
SOURCE	fruit fly,		
ORGANISM	Drosophila melanogaster		
AUTHORS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 797)		
TITLE	Genoscope.		
JOURNAL	Direct Submission		
COMMENT	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)		
	- Web : www.genoscope.cns.fr		
	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).		
	The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named Rpci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and Est libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .		
FEATURES	location/Qualifiers		
source	1..797		
	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"		
	/clone_lib="RPCI-98"		
	/clone="BAR08C03"		
	/note="end : TET3"		
BASE COUNT	170 a 78 c 89 g 133 t 327 others		
ORIGIN			
Query Match	2.8%; Score 44; DB 122; Length 797;		
Best Local Similarity	13.6%; Pred. No. 0.32;		
Matches	54; Conservative 176; Mismatches 167; Indels 0; Gaps 0;		
QY	603	ACAAAACCCACAATTTTATTGAGAGCTAGAAAAGTTAGGCTGACAACTTACTGG	662
Db	398	WBMMAAAYVATSSVASMSTSHHWVDTWTWKKAHVHKKMKDKKKVMMMAADRSK	457
QY	663	TCATGTTCTTCAGTTTAGAAGATTGGTTCAGTTAGTATCAAAAGTTAAATATGATAT	722
Db	458	NAKKWKMDAAAAAANKKDRRAKKRKKSMKKDVAVMATKKAAMVVKNNAKKNNKK	517
QY	723	TTCAATTAATATGATCCAGAAAAGATTGGAGGTTTGGCATAGTATATCTAG	782
Db	518	NNNNNNKKRAAATKAKSNAKKDMSNATBHRTDAKKBHKSBRADRDKRYSVAAAKK	577
QY	783	TCGATATGCAATGCTCAAGAAATCTCAGTTATTTCTACCCGAGATACCATTCCTAA	842
Db	578	GKGKRAVYNNKVBISMAATRAWCKSVAAAMKRRSSNIBBKDDTDDVHKVHDANAA	637
QY	843	AGCGGCTATTATGCAACCACTGCAAAATGGCATCAAGCGAGTAAAGAGTCCCTTGA	902

Db 638 DNVAMVHKKAAAMVMAAMAAAAAMKMAAMAMAMMTBKKDKDKKAMCM 697

QY 903 ATCATGCCGGAATACAGATGGAATTCGGCGCTCATGCGAGTAATGATTCCTTT 962

Db 698 WBMCMCMMAAAMAAAMAAKSCMCATRRADCCATATYTRCHTKRDDMAAAMCMWYAYB 757

QY 963 ACCCGCGGATCGTATCATGATGATATTTGAAAGTG 999

Db 758 MMCCCYAAAYKNTWMTNMTDMHATKRMAMAKKD 794

RESULT 6

LOCUS A0940248 700 bp DNA GSS 27-JAN-2000

DEFINITION Sheared DNA-42E21.TF Sheared DNA Trypanosoma brucei genomic clone Sheared DNA-42E21, genomic survey sequence.

ACCESSION A0940248

VERSION A0940248.1 GI:6763513

KEYWORDS GSS.

SOURCE Trypanosoma brucei.

ORGANISM Trypanosoma brucei

REFERENCE 1 (bases 1 to 700)

AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C., Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S., Donelson,J., Fraser,C. and Adams,M.

TITLE Determination of clone end sequences from Trypanosoma brucei GUTat 10.1 sheared DNA library

JOURNAL Unpublished (1999)

COMMENT Other_GSSs: Sheared DNA-42E21.TR

Contact: Najib M. El-Sayed

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: nelsayed@tigr.org

Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared DNA library constructed at TIGR. Clones will be available for distribution through ATCC. Sheared DNA end sequences search page: <http://www.tigr.org/tdb/mbd/tbdb/>.

Seq primer: M13-Forward

Class: Shotgun.

FEATURES

source

Location/Qualifiers

1..700

/organism="Trypanosoma brucei"

/strain="TREU927/4 GUTat 10.1"

/db_xref="taxon:5691"

/clone_lib="Sheared DNA-42E21"

/clone_lib="Sheared DNA"

/note="Vector: pUC18; Site_1: SmaI; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (approx 2 kb). The v + 1 method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (making small insert libraries for whole genome shotgun sequencing projects. In genome sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

BASE COUNT 374 a 43 c 41 g 242 t

ORIGIN

Query Match 2.7%; Score 42.6; DB 117; Length 700;

Best Local Similarity 43.2%; Pred. No. 0.73;

Matches 201; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 915 TACACATGGCAATTCGGCGCTCATGCGAGTAATGATTCCTTTAACCGCGCATCG 974

Db 220 TGCACATTGTAATTAATGTCGACACATTAAAGTATTGTTGTAATAGCAGTAATA 279

QY 975 TATGATGATGATTTTGAAGGATGTTGATTCATGATCATGATCATGATGATGCGCG 1034

Db 280 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 339

QY 1035 TACAGATTGCTGAGAGAAATAGCTGAGCTTACCGCGGAATTAAAGATTATTCAGTTAT 1094

Db 340 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 399

QY 1095 TCAAGCCGAATTAATAGCATCTGCTAGTAGTGACCATTAATATTCATGATAATC 1154

Db 400 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 459

QY 1155 CATTAATCTCATGATTAATTAATTTATATGTTATACGATGAGAGATTTTAAAGCCAG 1214

Db 460 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 519

QY 1215 CGCAGACTACAAATTTCTCGAGAAATGCTCAAAACCCATTCAGTGATGAGCGA 1274

Db 520 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 579

QY 1275 GAAAAAATAGTCTCGATTAAGACCTTCTTGAGAGTGAATTAAGAAACCGGCGCTT 1334

Db 580 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 639

QY 1335 GGGTAATCTGAAAAACCATCTCTTAAATTAAGATTAATGA 1379

Db 640 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 684

RESULT 7

LOCUS A0038989 579 bp mRNA EST 29-MAR-1999

DEFINITION A0038989 Dictyostellium discoideum SS (H.Urushihara) Dictyostellium discoideum cDNA clone SSM246, mRNA sequence.

ACCESSION A0038989

VERSION A0038989.1 GI:3985742

KEYWORDS EST

SOURCE Dictyostellium discoideum.

ORGANISM Dictyostellium discoideum

REFERENCE 1 (bases 1 to 579)

AUTHORS Morio,T., Urushihara,H., Saito,T., Uegawa,Y., Mizuno,H., Yoshida,M., Yoshino,R., Mitra,B.N., Pi,M., Saito,T., Takemoto,K., Yasukawa,H., Williams,J., Maeda,M.N., Takeuchi,I., Ochiai,H. and Tanaka,Y.

TITLE The Dictyostellium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development

JOURNAL DNA Res. 5 (6), 335-340 (1998)

MEDLINE 99156227

COMMENT On Jun 5, 1998 this sequence version replaced gi:3189400.

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan

Email: d402huesakura.cc.tsukuba.ac.jp

PROJECT - Dictyostellium discoideum cDNA project in Japan.

FEATURES

source

Location/Qualifiers

1..579

/organism="Dictyostellium discoideum"

/strain="AX4"

/db_xref="taxon:44689"

/clone_lib="SSM246"

/clone_lib="Dictyostellium discoideum SS (H.Urushihara)"

/dev_stage="slug"

BASE COUNT 286 a 31 c 52 g 210 t

ORIGIN

Query Match 2.7%; Score 41.8; DB 48; Length 579;

Best Local Similarity 45.0%; Pred. No. 1.1;

Matches 157; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY	970	GATCGTATTCGATGATGATTAATTTTGAAAGCATGTGGTGAATCAATGCATCATGCTGT	1029
DB	71	GCTCATATTTTTTGATGAAAATATTATAATATTTCATAGTAATAAATAATTAATTAATGACAT	130
QY	1030	GCCCGTAGCAGAGTTGCGTGAGAATTAGCTGAGCTACCGCCGAATTAAGAATTATCCA	1089
DB	131	GAT	190
QY	1090	GTTATTCAGGCCGAATTAATTAAGATCTGCTAGTAGAGGACCATATAATTCATGAT	1149
DB	191	AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	250
QY	1150	AAATCCATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1209
DB	251	AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	310
QY	1210	GCCAGCGCAGAGTACACAAATTTCTCGAAGAAATGCTCAACACCATTCAGGTGATGG	1269
DB	311	TGGATTAACAACATGTAATGATATGATATCAATCAATATTTTCAATTAATTAATCAAGAGT	370
QY	1270	AGCGAGAAAAAATAGCTCCGATTAAGAGCTTTCTTGAGAGGAGAAATA	1318
DB	371	ATGAACAAGGTTGTTTATTAATAAATGAATATAAATGAATTAATGAATA	419
RESULT	8		
CNS0182P			
LOCUS			
DEFINITION	CNS0182P	1101 bp	DNA
	Drosophila melanogaster genome survey sequence SP6 end of BAC		GSS
	BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit		
	fly), genomic survey sequence.		
ACCESSION	AL108811		
VERSION	ALI08811.1	GI:5629115	
KEYWORDS	GSS.		
SOURCE			
ORGANISM			
	fruit fly.		
	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1101)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CBPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pbeloBAC11. Location/Qualifiers 1..1101 /organism="Drosophila melanogaster" /plasmid="pbeloBAC11" /db_xref="taxon:7127" /clone_1fb="DrosBAC" /clone="BACN37D10" /note="end : SP6"		
COMMENT			
FEATURES			
Source			
BASE COUNT	274 a	268 c	128 g
ORIGIN			
	73 t	358 others	
Query Match	2.7%	Score 41.8;	DB 123; Length 1101;
Best Local Similarity	17.1%;	Pred. NO. 1.4;	
Matches	62;	Conservative 164;	Mismatches 133;
		Indels 4;	Gaps 1;
QY	979	GATGATGATATTTGAAAGTGTGTTGATTAAGAATCAATCATGTCATGCCGTAGC	1038
DB	742	GKMDGTATAMTWTWATMTWKAAITDDAKPAAGERRRDKRKRTARADGGRRARTFRFRAW	801

QY	1039	AAAGTGCCTGAGATTGCGAGCTTGCAGCCCGCAATTAAGATTATCGATTATCA	1098
Db	802	AAAGGARAAGARRARAARRADDDEHMDMAAAAAAAAAAAATTTWDRWMDMDWR	861
QY	1099	GCCGAATTAATAAGCATCTCTAGTAGTGGCACCATTAATTCATGATTAATTCATT	1158
Db	862	D----DPTTAAMWDDARARARRRRRRRRRRARRAARRADDPTDRMADATTDKDT	917
QY	1159	AATCGCATGATTAATAATTTATGTTATACAGATGAAGAGATTTTAAAGCCAGCA	1218
Db	918	KTTWTDDEDDMKARDKRWAAKADGAMKWRPARDMDATKADDGWDKMGGRKRGK	977
QY	1219	GAGTACAAATATTCGAGAAAAATGCTTAACACCATTCAGTGGATGGAGGAGAAA	1278
Db	978	DKRRDKRTGKTGKKDDDDDKXTWTRNDMMWMTRTKDWMDWDDGRGRGRTKRKMGAMWRAD	1037
QY	1279	AAATATAGTCGTGATTAAGACCTTCTTGGAAGTGAATAAAGAACCCGGGCGTTGGCT	1338
Db	1038	AAARDDTDGKDTFTADKRRKRTDFTTKRGGDDMRKDRKRRKDDKTKKDATWDDDD	1097
QY	1339	AAT 1341	
Db	1098	ARD 1100	
RESULT 9			
CNS003F8/c	LOCUS		
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #		
ACCESSION	BACR08C03 of RPCI-98 library from Drosophila melanogaster (fruit		
VERSION	fly), genomic survey sequence.		
KEYWORDS	AL064634		
SOURCE	GSS.		
ORGANISM	fruit fly.		
REFERENCE	Drosophila melanogaster		
AUTHORS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
JOURNAL	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
	1 (bases 1 to 797)		
	Genoscope.		
	Direct Submission		
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :		
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)		
	Determination of this BAC-end sequence was carried out as part of a		
	collaboration with the Berkeley Drosophila Genome Project (BDGP).		
	The BDGP is constructing a physical map of the Drosophila		
	melanogaster genome using these BACs. For further information		
	please see http://www.fruitfly.org The BDGP Drosophila		
	melanogaster BAC library was prepared by Kazuhiro Oosagawa and		
	Aaron Mammoser in Pieter de Jong's laboratory in the Department of		
	Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,		
	NY. The library is named RPCI-98 and was constructed by partial		
	EcoRI digestion of Drosophila DNA provided by the BDGP from the		
	isogenic strain Y2; cn bw sp, the same strain used for the BDGP's		
	P1 and EST libraries. A more detailed description of the library		
	and how to order individual BAC clones, the entire library, or		
	filters for hybridization from the BACPAC Resource Center can be		
	found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .		
	Location/Qualifiers		
FEATURES	1..797		
source	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"		
	/clone_lib="RPCI-98"		
	/clone="BACR08C03"		
	/note="end : TET3"		
BASE COUNT	170 a	78 c	89 g 133 t 327 others
ORIGIN			
Query Match	2.7%; Score 41.6; DB 122; Length 797;		

THIS PAGE BLANK (USPTO)
